COGNITIVE PSYCHOLOGY 21, 469-514 (1989)

Dissociations and Dependencies between Speed and Accuracy: Evidence for a Two-Component Theory of Divided Attention in Simple Tasks

HAROLD PASHLER

University of California, San Diego

Previous work has amply demonstrated divided attention "costs" both in single-task multistimulus visual processing and in performance of multiple simple *tasks;* however, the relationship between the two has not been clarified. This article postulates two distinct causes: (1) visual processes that commence without delays and proceed simultaneously, but show lingering mutual interference dependent upon complexity, and (2) discrete queueing of the response selection stage. The first has resource-like properties, while the second has bottleneck-like properties. Either or both can generate performance costs observed in any particular situation, accounting for a variety of previous results. To test this theory, the effects of stimulus onset asynchrony (SOA) on accuracy and speed in performing dual choice tasks were examined. The first two experiments involved a choice response to a tone as a first task, and a second task requiring complex perceptual decisions (digit identification or conjunction search) with masked displays and *unspeeded* second responses. Reducing the SOA had negligible effects upon second-task accuracy, and performance in the two tasks was virtually independent. However, when speeded manual (Exp. 3) or vocal (Exp. 4) responses were required on the same second task, dramatic interference was observed, with strong positive dependencies between reaction times (RTs) on the two tasks. When both tasks involved complex visual displays, SOA reductions produced dramatic interference, but no dependencies between performance, whether the first task involved a speeded (Exp. 5) or unspeeded (Exp. 6) response. The results reject pure late-selection accounts and general capacity sharing models, and support the two-component theory. They also suggest that standard use of the term "attention/' suggesting a single resource or mechanism, is highly misleading.

The author is grateful to Gordon Baylis, Jim Hoffman, James Johnston, Jeff Miller, Don Norman, and Allen Osman for useful discussions of this work and related issues, and also to Melissa Berhow, Melissa Darnell, Alex Fleishman, Paul Havig, and Shihyen Hsu for excellent assistance in conducting the experiments reported here. Geoffrey Loftus and two anonymous reviewers made useful comments on an earlier version. Support for this work was provided by NASA under Interchange NCA2-225.
Correspondence should be addressed to the author at the Department of Psychology, C-009, University of California San Diego, La Jolla, CA 92093.

0010-0285/89 $7.50

INTRODUCTION

Basic Goals

People encounter severe limitations when they attempt to perform more than one task at the same time, even when the tasks are very simple. Understanding these limitations should have important theoretical and practical implications. Theoretically, such understanding should provide insights into the nature and control of elementary mental processes, while on the practical side, an understanding of dual-task limitations should be useful in the rational design of complex systems that require human-machine interaction. The goal of the present article is to advance a concrete account of how the dual-task interference effects observed when subjects attempt to perform two simple tasks simultaneously are related to capacity limitations arising in perceptual processing of complex visual displays. These two aspects of divided attention have traditionally been investigated separately. Perceptual processing limitations have been revealed in situations in which multiple stimuli must be processed in a single coherent task. Mostly, such studies involve a single response that can only be selected once the subject has processed many distinct, simultaneously presented stimuli, e.g., visual search. On the other hand, limitations evident when two separate tasks must be performed are studied by combining two tasks, each involving an unrelated mapping of a set of possible stimuli to a set of possible responses (e.g., Welford, 1958; McLeod, 1978).

In each of these two kinds of situations, there is clear evidence of divided attention costs. However, little has been done to characterize the relationship between the divided attention limitations observed in these two kinds of research. Both research and theory are mostly bound to one sort of paradigm or another. In many discussions, the term "attention" is used to refer to a limited capacity allocated in the visual field, and also to refer to limitations responsible for interference between tasks. Thus, following ordinary usage, many writers speak of both "devoting attention to stimuli" and "devoting attention to tasks." This article argues that capacity limits arising in perceptual processing of multiple objects are quite separate from the postponement of central cognitive operations of decision and response selection that occurs when multiple responses must be selected. The two limitations are not only separate, but also fundamentally different from each other in character. According to the theory proposed here, impairment of performance in any particular divided attention situation may be due to the occurrence of either or both of these forms of interference. If the account advanced here is correct, then the common use of the term "attention" to cover all of these various limitations is

probably very misleading, as suggested by Broadbent (1982). The suggestion that dual-task interference may not be a homogeneous phenomenon is hardly novel, having been advocated in the form of theories postulating multiple mental "resources" (e.g., Gopher, Brickner, & Navon, 1982; Wickens, 1980). However, both the conclusions reached here, and the methods used to reach these conclusions, are very different from those of the multiple resource approach; these differences will be discussed more fully in the General Discussion.

We begin with a selective review of some important empirical generalizations concerning costs of divided attention. The first-section reviews evidence pertaining to simultaneous performance of multiple tasks, evidence that makes a strong case for postponement of central decision and/or response selection as the fundamental cause. The second section reviews experiments involving numerous and/or complex visual discriminations, and concludes that capacity limitations arise here that are clearly attributable to the perceptual processing stage. Previous hypotheses about how these two sorts of limitations might be related are then discussed, and a tentative theory is advanced on the basis of the findings reviewed. This theory is then tested in six experiments which reveal dissociations in the effects of temporal overlap on different response measures, and examine patterns of dependencies between performance on two overlapping tasks. These results confirm the response selection postponement account and demonstrate the independence between this postponement and capacity limits on perceptual processing.

Interference between Discrete Tasks

The Psychological Refractory Period Paradigm

The first question is what happens when people try to perform two simple *tasks* at once, each involving a separate response based upon a distinct stimulus. The simplest and most straightforward paradigm that requires this is the psychological refractory period (PRP) paradigm. In the PRP paradigm, two stimuli (S1 and S2) are presented in rapid succession, separated by a stimulus onset asynchrony (SOA); the subject must make a response to each (R1 and R2, respectively) as rapidly as possible. The phenomena that arise here were extensively investigated in the 1960s, and various accounts were proposed (for reviews, see Bertelson, 1966; Smith, 1967). The most basic observation here is a slowing of the second response, more so as the SOA is shortened. With even very simple tasks (say, two two-choice tasks with simple stimuli and manual responses), this interference is very sizable, generally involving delays of at least several hundred milliseconds.

Accounts of the Phenomenon

Two types of models have generally been considered to account for multiple-task interference: *postponement models* and *capacity sharing models.* Postponement models propose that specific cognitive operations can only occur when a single mechanism is *exclusively* dedicated to performing that operation for a sufficient period of time. When the critical mechanism is occupied with one task, processing operations in the other task that require the mechanism must be postponed until the mechanism becomes available; hence the concept of a processing "bottleneck" or "single channel." Different models locate the postponement in different processing stage(s). Most frequently discussed are (1) *perceptual identification,* (2) *decision and response selection,* and (3) *response initiation and execution* postponement models. Broadbent's original filter theory of attention (Broadbent, 1958) proposed that identification of stimuli was the bottleneck responsible. Welford (1952, 1980) and Smith (1967) proposed that decision and response selection processes constitute the bottleneck. More recently, it has commonly been argued that in simple tasks the primary source of interference does not arise until the initiation or execution of motor responses (Keele, 1973; Logan & Burkell, 1986; Norman & Shallice, 1985).

The major alternative class of models to postponement models are capacity or resource theories (Kahneman, 1973; Norman & Bobrow, 1975; Wickens, 1983). These accounts propose that interference between tasks originates not in postponement of particular stages or operations, but rather in a reduction of the efficiency with which each task operates (albeit simultaneously)—caused by a graded sharing of resources between the tasks. The simplest version of this type of theory supposes that just a single very general resource is allocated to support all cognitive processes (Kahneman, 1973) and McLeod (1977b) applied this kind of model specifically to the PRP paradigm.

"Classic" Studies of PRP Effects

The early work on the PRP paradigm provided some interesting results, but these did not prove to be diagnostic with respect to the theories just described. According to the postponement model, some bottleneck stage(s) cannot operate simultaneously for each of the two overlapping tasks. This hypothesis yields the apparently straightforward prediction that as the stimulus onset asynchrony between S1 and S2 is reduced, reduction in the SOA beyond a certain point produces a corresponding increase in the duration of R2 (measured front S2). The degree to which this minus-one slope prediction is borne out has been debated (see Bertelson, 1966; Kahneman, 1973; Kantowitz, 1974). It does seem to fit

most data quite well, however, for overlapping choice-reactions.[1] The minus-one slope prediction is not very diagnostic, however, since it may also be consistent with versions of capacity sharing models (cf. McLeod, 1977b). In addition, it cannot discriminate between different postponement models. It might seem that the postponement model should also predict that at relatively long SOAs, where the bottleneck stages in the two tasks would not overlap, dual-task R2 should be as short as the same R2 made in a single task alone. This is not usually observed. However, a general cost of doing two tasks at once might arise because two tasks cannot be *prepared* as well as one; this would be perfectly consistent with postponement accounts (Gottsdanker, 1980; Pashler, 1984a).

Another finding that has frequently been observed with the PRP paradigm is a slowing of the first response (R1) in the first task (Gottsdanker & Way, 1966; Herman & Kantowitz, 1970; Kahneman, 1973). R1 slowing is readily accounted for with capacity models, since both tasks are assumed to be performed with depleted allocations of capacity. Postponement models need to be supplemented if they are to account for R1 slowing. The preparation factor noted above is one likely contributor. In addition, R1 slowing may result from a strategy called "grouping." Unfortunately, this term has several different meanings, all of which share the core idea that the two responses are coupled in some way. Borger (1963) suggested one possible grouping strategy: the first response is saved until the second response has been selected, so that the two responses can be emitted in rapid succession. Thus, a dramatic slowing of R1 would result, whether or not the second task interferes directly with the first task. In this way, R1 slowing could be accounted for without discarding postponement models, and thus R1 slowing, by itself, is not especially useful in discriminating between theories.

Recent Analyses of the PRP Paradigm

The basic results observed in the early work with overlapping tasks seem broadly consistent with a wide variety of causal accounts for dual-task slowing. A more diagnostic way of testing postponement models, and empirically distinguishing them from capacity sharing models, was developed by Pashler (1984b) and Pashler & Johnston (1989). The method relies on manipulating stimulus factors to increase the duration of selected processing stages. How can this distinguish between different models? The logic is shown in Fig. 1 which illustrates the response selection postpone-

[1] Except that the function relating R2 to SOA rarely shows a sharp elbow; this is to be expected, given that the durations of all component states undoubtedly have substantial variability.

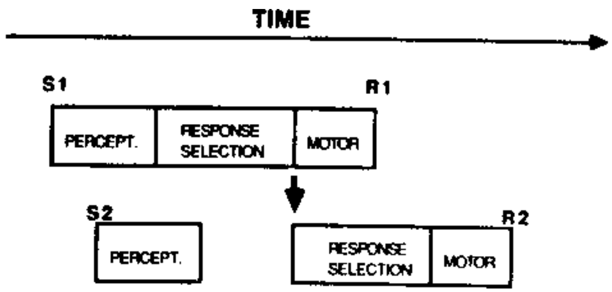

FIG. 1. The sequence of stages in response selection postponement: response selection on task 2 waits for the completion of response selection on task 1, while the other stages can overlap.

ment hypothesis. Response selection in the second task does not begin until response selection in the first task is complete, whereas the perceptual processing of S2 begins as soon as it is presented.

Pashler's (1984) method for testing postponement models utilizes experimental S2 variables that selectively slow either perceptual or response selection stages of the second task. Consider a factor that slows response selection, and compare its effect on R2 latencies in (a) the dual-task condition shown in the figure, and (b) a single-task control, not shown. Such a factor will slow R2 the same amount in both dual-task and single-task conditions. Stated generally: If a factor slows down stages of processing located at or beyond the locus of any single-channel bottleneck, then its effects will be *additive* with the dual-task vs. single-task slowing.

Consider, on the other hand, a factor *slowing perceptual processing* of S2. In a single-task control, the entire factor effect appears as a slowing of the observed reaction time (RT). In the dual-task condition, however, response selection on the second task cannot begin until both its input and the processor are available. Therefore, the factor effect will be partially or completely "washed out": on trials where response selection waits on the processor and not the input, lengthening the perceptual processing time for S2 will have no effect. As SOA is reduced, the probability of this happening will approach unity. Stated generally: If a factor slows down stages of processing prior to the locus of the postponement, its effects will be *underadditive* with the dual-tasks vs. single-task slowing, and underadditive with SOA in the dual-task condition.

One advantage of this method is that it allows nonobvious predictions to be derived from postponement models. Overadditive interactions in latencies are much more common than underadditive interactions (and also follow straightforwardly from capacity theories; cf. McLeod, 1977b); this makes underadditivity predictions especially diagnostic for testing

postponement models.[2] This analysis is a special case of PERT networks, described and applied in important papers by Schweickert (1978).

A number of experiments along these lines have been reported, using pairs of choice reaction time tasks (Pashler, 1984a; Pashler & Johnston, 1989). The results strongly favored response selection postponement models. In several experiments, the effects of the intensity of a visual S2 were greatly reduced in a dual-task condition, more so as SOA was shortened. On the other hand, factors affecting response selection—decision outcome in visual search (Pashler, 1984), and intertrial repetition in a choice-reaction time task (Pashler & Johnston, 1989)—had effects additive with dual-task slowing. These results are inconsistent with alternative postponement models, and also with general capacity sharing models. As noted earlier, the only strong empirical support for capacity sharing models of simple overlapping tasks comes from the phenomenon of R1 slowing. Postponement models can be reconciled with this slowing if the first response is sometimes grouped with the second. Extending the methods just described, Pashler & Johnston (1989) tested the grouping account of R1 slowing, by examining factor effects and interactions when R1 slowing was either encouraged or discouraged. Detailed support for the grouping account was obtained, thus undermining the original support for capacity sharing models.

This brief review of simple dual-task studies suggests no role for perceptual capacity limits. However, the stimuli used in virtually all of the experiments described require only the simplest discriminations, e.g., two or three choices involving just a single tone or letter. We turn now to the second type of divided attention research, where only a single coherent task is performed, but many complex perceptual discriminations may be required.

Interference between Processing of Multiple Stimuli

Another important tradition of experimental work has examined the limitations that arise when different stimuli must be processed, with a single response somehow made contingent upon all of the stimuli. The prototypical example is visual search, where the contingency is disjunctive ("respond 'yes' if any of the items in the display is a target"). This task was introduced to look at limitations in perceptual processing without placing memory demands such as whole report (Neisser, Novick, & Lazar, 1963). Many studies examine impairment in visual search accuracy

[2] Note that relaxation of the idea of the strict successiveness of the stages (McClelland, 1979) would lead only to a quantitative, not qualitative, change in the predictions. In any case, empirical evidence currently favors a discrete transition between stimulus evaluation and response selection (Meyer, Yantis, Osman, & Smith, 1984; Miller, 1988).

and RT as the number of stimuli ("display size") is increased. With single feature targets, increases in display size generally produce negligible increases in RTs (Green & Anderson, 1956; Treisman & Gelade, 1980). By contrast, when subjects search for an alphanumeric character among distracting characters, each additional item usually costs the subject from 10 to 60 ms (see Gagnon, Cavanagh, & Laurencelle, 1978, for a review). Slopes are reduced when the discriminability of the target and background elements is increased (e.g., Duncan, 1983; Pashler, 1987b), or by consistent practice (e.g., Kristofferson, 1972), or by homogeneity of the background elements (e.g., Gordon, 1968). Despite these useful empirical generalizations, the source of these effects is not entirely clear, largely because many different factors could potentially contribute (for a clear review, see Duncan, 1980b). Such effects are often attributed to serial processing (e.g., Atkinson, Holmgren, & Juola, 1972; Treisman & Gelade, 1989). However, they might also be due to capacity limits on parallel processing (Rumelhart, 1970; van der Heijden, 1975). Display size functions alone cannot distinguish these accounts (Townsend & Ashby, 1983).

In the case of search involving small displays of alphanumeric characters, however, several lines of evidence argue against serial identification of items in the display. On the one hand, effects of display size are additive with the effects of visual degradation of items in the display (Pashler & Badgio, 1985). If the perceptual processes affected by degradation were operating sequentially, multiplicative interactions should have occurred. Another approach, developed by Eriksen and Spencer (1969) and Shiffrin and Gardner (1972), compares accuracy in search of briefly presented displays, as a function of whether all the items are presented simultaneously, or instead, portions of the display are presented sequentially. Serial or indeed any capacity limited models clearly predict a major advantage for the sequential presentations, yet such advantages are not observed.

More recent work, however, shows that as the discrimination difficulty or display size is increased, this equality ceases to hold. Thus, Kleiss and Lane (1986) found large successive advantages with target/distractor sets requiring fine discriminations, but not with a set that apparently did not. Similarly, Badgio and Pashler (1988) found that both increasing display size and decreasing discriminability produced successive advantages (see also Prinzmetal and Banks, 1983). Such effects are not attributable to decision noise or other statistical effects, and therefore they reject the claims of Shiffrin and Gardner (1972), Duncan (1980a), and other theorists.

In summary, divided attention costs appear in tasks requiring a single response based on the identity of multiple stimuli, when large complex displays are used. These capacity limits might indicate serial scanning, or some form of capacity limits on parallel visual processing.

The Relationship between Multistimulus Effects and Multitask Effects

The preceding sections have briefly reviewed evidence for divided attention costs in two sorts of situations. In the first, multiple tasks must be performed simultaneously. With even the simplest task combinations, delays in the hundreds of milliseconds are common. Second, in tasks that require subjects to process multiple stimuli but generate only a single response, divided attention costs sometimes appear. With displays requiring sufficiently numerous and complex discriminations, one finds delays in the 10 to 50 ms range on speed of responding, for each additional display item. Effects on accuracy are more difficult to estimate (see Dun-can 1980b), but are often sizable: Badgio and Pashler (1988) found that simultaneous presentation sometimes reduced two-alternative forced choice accuracy by 10%, compared to successive presentation.

We turn now to the basic focus of this article: What is the connection between the decision/response selection postponement phenomenon, argued for above, and the perceptual capacity limits just described? Although most investigators have not attempted to relate these two classes of phenomena, some unified interpretations have been proposed, explicitly or implicitly. We start by mentioning two much-discussed hypotheses that *cannot* be sustained in light of the evidence reviewed above.

1. Late-selection theory. One important view of attention is that of ("late-selection") theorists who have argued that perceptual processing and object identification are fully automatic. This position holds such processes are not subject to capacity limits, or to any form of attentional control. Shiffrin and Gardner (1972) and Duncan (1980a) provide clearly argued statements of this view, maintaining that *all* interference effects are postperceptual, and that what *appear* to be perceptual limits prove, on closer examination, to arise later in processing. Thus, Duncan (1980a) proposed that a "transfer mechanism" is necessary not only for decision and/or response selection, but also for short-term memory preservation with briefly presented stimuli. Similarly, Shiffrin and Gardner suggest that "the entire processing system prior to short-term memory" is free of capacity limits (p. 81). Such accounts would be quite consistent with response selection postponement in multiple tasks. However, they can no longer be defended in their strong form, given the clear demonstrations of capacity limits provided by the work of Kleiss and Lane (1986), Prinzmetal (1983), and Badgio and Pashler (1988). (The same applies to the work of Posner and Boies, in the well-known papers which concluded that perceptual processing "does not take capacity").

2. General capacity theories. While late-selection theorists have denied the existence of perceptual capacity limits, general capacity theories (Kahneman, 1973; McLeod, 1977b) have proposed that divided attention effects stem from a common pool of resources used to "power" cognitive operations. If performance is monotonically related to the capacity available to a task, then simultaneous performance will produce impairments. On this account, multistimulus and multitask costs have the same source—namely, reduction in the capacity available to component stages or processes. This view is suggested, albeit less explicitly, in other theoretical discussions. For instance, Broadbent (1982) says that in dual tasks, "the processes occurring within the person are numerous and widespread physically; but the empirical data on task combination suggest that each feature extracted from the sensory buffer affects those processes very widely" (p. 286). The key idea here is that the division of resources is *graded,* and that the effects of depletion apply to all stages. The reader will recall, however, that the chronometric work described earlier concerning multitask interference (Pashler, 1984; Pashler & Johnston, 1989) provides no support for this hypothesis as applied to interference between simple tasks: instead, discrete postponement of processing of central response selection stages seems to cause delays when two tasks are performed simultaneously.

Thus, neither late-selection theory nor general capacity theory, as usually formulated, provide a satisfactory account of the divided attention effects reviewed above. Late-selection accounts fail because they do not predict any interference at the perceptual stage. General capacity accounts attribute interference between simple tasks to graded sharing of general resources; this is refuted by the evidence cited earlier, indicating discrete postponement of central stages in multiple tasks.

However, it will probably have occurred to the reader that these considerations do not necessarily rule out all accounts that attribute multitask interference and multistimulus performance decrements to a common cause. Suppose, for instance, that a single mechanism is always involved whenever a response must be selected, and when this device is selecting one response, it cannot select another—hence, response selection postponement in the PRP paradigm. However, this device may *also* be required for perceptual processing *under certain circumstances*—namely, when the perceptual processing is sufficiently demanding, e.g., with difficult and numerous discriminations between target and distractor elements in visual search. This account would predict parallel capacity-free processing with small, simple displays (e.g., Shiffrin & Gardner, 1972), but not with larger, more complex displays (e.g., Kleiss and Lane, 1986). This single attentional mechanism might be directly involved in carrying out particularly demanding search tasks like detection of feature conjunctions, claimed by Treisman and Gelade (1980) to operate sequentially. We might somewhat whimsically term this view the "all purpose central processor unit (CPU)" theory. Something like this view seems to have been suggested by Posner (1982).

It might seem that comparing the sheer magnitude of different divided attention delays could refute this theory, but it cannot. For instance, a 200-ms PRP effect might reflect the CPU being occupied for hundreds of milliseconds in selecting a response in a first task, while the 30 ms/item slope in conjunction search might reflect rapid scanning carried out by this same device. Plainly, the CPU might take longer to do some things than to do others. Can any of the PRP results sketched earlier falsify this account? The effects observed by Pashler (1984b) and Pashler and Johnston (1989) argued against postponement of perceptual processing, for reasons indicated above; however, this too is consistent with the all purpose CPU account, since second-task visual discriminations in those experiments were very simple, probably less complex than those used by Duncan (1980a) or Shiffrin and Gardner (1972).

Nonetheless, several considerations speak against the CPU account. When capacity limitations do arise in perceptual processing, they seem not to reflect serial processing at all. Badgio and Pashler (1988) observed superior performance in successive compared to simultaneous displays, but not until the interval between the two successive displays was lengthened to at least several hundred milliseconds. Under conditions in which successive displays were processed more accurately than simultaneous displays, performance was unimpaired when the presentation order of the two-component displays was completely unpredictable to the subject (see also Pashler & Badgio, 1987). It seems implausible that sequential scanning by a single mechanism could fail to benefit when the order of scanning was known in advance. By contrast, substantial effects of predictability of stimulus order are found in multitask experiments. Pashler (1988) found that when a PRP task was altered so that order of stimuli was unpredictable to the subject, response time delays were greatly magnified (especially when both responses were manual).

These differences make it appear unlikely that perceptual capacity limitations (e.g., Kleiss & Lane, 1986) and the response selection postponement phenomenon in dual-task situations (e.g., Pashler & Johnston, 1989) reflect a single underlying attentional mechanism that must be deployed for both central decisions and complex perceptual processes. The CPU account is not compellingly refuted, but neither is it supported by what data does exist. For this reason, therefore, an alternate account is sketched that can fit these various data, and then more incisive empirical tests are undertaken.

A Two-Component Theory

I propose that "divided attention costs" can arise in two ways. The first pertains to selection of responses (and perhaps to complex cognitive operations in some more general sense—a matter for further research). Whenever a response must be selected, a single mechanism must carry out this job.[3] When this mechanism is servicing one task, it is not available to work on another task. Response selection proceeds most efficiently when this mechanism can be preset for the first task that will arrive, the preparation factor noted above. The second source of divided attention costs pertains to perceptual processing. (Existing evidence deals purely with the visual modality, so the discussion is restricted to that case.) This second component, unlike the first, *does* involve something that amounts *roughly* to a graded allocation of resources, rather than discrete serial processing. When new stimuli arrive, they can "grab" the mechanism without the system being prepared in advance (Badgio & Pashler, 1988). Indeed, the many well-known effects of unattended stimuli (e.g., Eriksen & Hoffman, 1973) illustrate that this sometimes occurs involuntarily. However, the resource metaphor is only approximate; Badgio and Pashler (1988) propose that mutual interference generated by cross-talk may be the underlying cause.

This overall conception of the two distinct components of divided attention costs is illustrated in Fig. 2. The figure shows the processing occurring when two stimuli arrive and generate interference of *both* types. As soon as each stimulus arrives, its perceptual processing commences immediately. If the total difficulty of the perceptual processing exceeds a certain level (represented by the width of the channel on the top of the figure), interference occurs, impairing the accuracy with which *both* stimuli are processed. As soon as the first stimulus is processed to criterion, the central mechanism begins selecting the appropriate response. When the second stimulus is encoded before the response selection on the first task is complete, postponement occurs (generating underadditive interactions with early stage stimulus factors, as reported by Pashler (1984b) and Pashler and Johnston (1989), reviewed above).

This account can readily explain the results described earlier. With

[3] Obviously speaking of a single mechanism here has purely functional significance; a neural network with patterns of activity that needed to be "flushed" between operations might qualify.

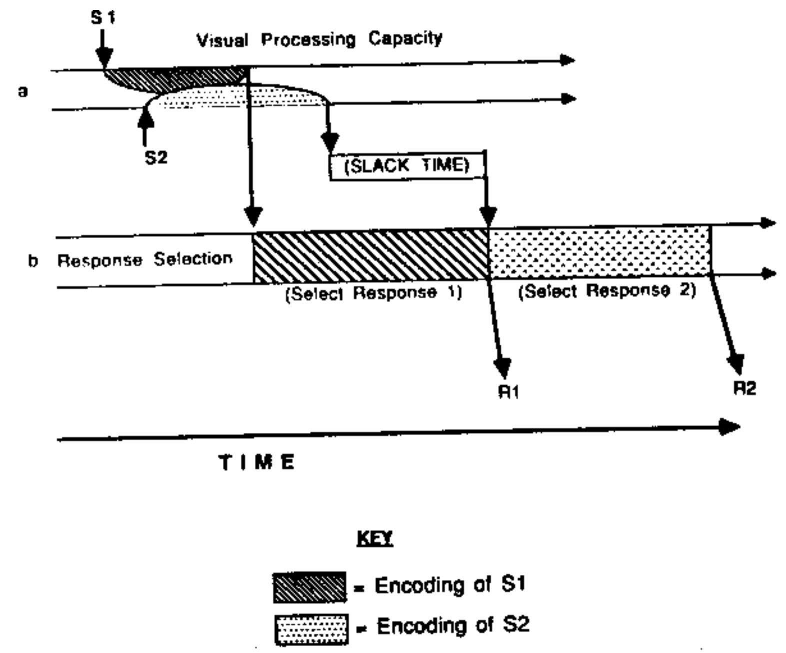

FIG. 2. The two-component model, illustrating interference at two points in processing: simultaneous degradation of perceptual processing, and queueing at response selection.

visual search tasks, only *one* response selection process occurs, even when there are many stimuli. Thus, when multiple stimuli redundantly indicate the same response in visual search, they apparently activate that response simultaneously (Miller, 1982), rather than queueing as in the PRP paradigm. Therefore, interference in visual search is entirely attributable to the perceptual component, as in the results of Kleiss and Lane (1986) and Badgio and Pashler (1988). In the PRP task, however, experiments have usually used very simple stimuli, often in different perceptual modalities, and thus only the second (response selection queueing) component is operative there.

The Present Experiments

The empirical work reviewed earlier is consistent with the two-component theory. The goal of the empirical work presented here is to provide much stronger tests of the two-component model, pitting it against alternative models. The theory states that whenever two simple tasks are performed with stimuli in close succession, selection of the second response cannot begin until the corresponding stage(s) of the first task are completed. However, perceptual processing of the second stimulus need not wait for any stages of the first task, and will only suffer interference from this perceptual processing when the total complexity of the perceptual demands exceeds some threshold. The chronometric studies reviewed earlier tested the prediction that when only response selection postponement is at work, slowing the perceptual processing of S2 will not greatly delay the second response, since postperceptual stages of that task must still wait for selection of R1. The studies reported here rest on another straightforward idea: if perceptual processing of S2 operates without interference, then the probability of it being successfully completed within a fixed time from when S2 became available will not be affected by whether the first task overlaps it in time. This can be arranged with a backward mask that terminates perceptual processing of S2 after a fixed interval from its onset.[4] The prediction, then, is that subjects should be no less accurate in identifying a brief masked S2 when S2 is presented at a short SOA after the first stimulus. At the same time, if the second task were speeded, and the masks were eliminated, a dramatic slowing of R2 would be expected as SOA is shortened.

The two-component theory thus predicts a dissociation between response measures on the effects of a single variable, namely SOA. With the very same task combinations, SOA should have dramatic effects on R2 latencies when R2 is speeded and S2 is unmasked, while it should have negligible effects on R2 accuracy when S2 is masked. With unspeeded R2s, the theory also predicts no dependencies between performance achieved on the two tasks on any given trial. Note that the method examines the effects of SOA, rather than comparing single-task and dual-task performances. Such comparisons are suspect because in a single-task condition the subject can *prepare* exclusively for one task, while effort and time devoted to preparation must be shared between tasks in the dual-task condition (on this point, see Gottsdanker, 1980; Logan, 1978; Pashler, 1984a). By contrast, manipulating SOA over a wide range reveals dual-task interference with preparation effects "partialled out."

The reasoning proposed here is analogous to that of Santee and Egeth (1982), who studied the interference generated by response-incompatible flanking characters in a two-choice task (Eriksen & Hoffman, 1973). San-tee and Egeth hypothesized that if flankers affect response selection, but not perceptual processing, then they should affect speeded RTs, but not accuracy with masked targets. This prediction was confirmed. The current approach is also closely related to the work of Salthouse (1981), who examined various stimulus factors under both accuracy and RT conditions, and found converging evidence for chronometric stage analyses. S-R compatibility, for instance, is widely supposed to affect the duration of response selection (Sternberg, 1969), and this factor had sizable effects on the function relating accuracy to speed under speeded instructions, but negligible effects on the function relating accuracy to mask delay. This nicely confirms that S-R compatibility affects the duration of stages that *follow* those terminated by the mask.

[4] The reader will note, as the paper proceeds, that no conclusions will rely on the assumption that the masks terminate perceptual processing instantaneously.

Dual-task studies in some ways similar to those proposed were reported by Blake and Fox (1969), who observed no reduction in accuracy for discriminating a single letter presented at threshold when the letter was presented after a tone requiring a speeded detection response. Unfortunately, those results are inconclusive, since the letter was not masked, and thus could have remained available in iconic memory until completion of the first task. In addition, the use of simple RT as the first task may eliminate the need for response selection, obviating central interference. In Experiment 1 below, the first task required a speeded choice response, and the second task required a "highest digit judgment" from a masked display of digits (Pashler & Badgio, 1985, 1987). Here, subjects report the numerically highest of these digits. Thus, the criterion for report is abstract, an element that is a target on one trial may be a distractor on another trial, and processing must be exhaustive to be assured of a correct answer. The task is actually surprisingly easy (e.g., compared to deciding whether there are two copies of any of the digits in a display). Pashler & Badgio used the task to argue that multiple digits could be identified in parallel—a conclusion that is independent of the dual-task question at issue here.

EXPERIMENT 1

The first experiment addressed the question of whether the accuracy in the highest digit task with masked displays would be impaired by overlapping this task with a first task requiring a rapid choice response to tone stimuli. In this experiment, subjects made a speeded two-choice response to a tone, which was high or low in pitch. The tone was followed at an SOA of 50, 150, or 650 ms by a brief array of eight digits, followed by a mask. (The 650-ms SOA is longer than the average R1 response time; the function of this condition is to make subjects prepare both tasks, but allow them to be performed with little overlap.) The subject had to determine what was the highest digit in this array (from among four choices), and make the appropriate *nonspeeded* response.

Because the first task does not involve complex visual processing, perceptual capacity limits should not appear. The first task is speeded, so the single-response selector mechanism will be occupied with selecting the first response as soon as the first stimulus has been perceptually processed. Perceptual processing on the second task should proceed without any interference on this account, however. Postponement of second-task response selection will occur, but it should not affect the accuracy measure employed here, just as incompatible flankers did not affect accuracy in Santee and Egeth's (1982) study.

Method

Subjects. Thirty undergraduates at the University of California, San Diego, participated as subjects in the experiment, in partial fulfillment of a course requirement.

Apparatus and stimuli. The stimuli were presented on Princeton Graphics SR-12 monitors, controlled by IBM PC microcomputers (equipped with Sigma Design Color-400 boards, providing a display resolution of 640 x 400 pixels). The first stimulus was a tone presented through the speakers on the monitors, at 300 or 900 Hz. The second stimulus was a centrally located display of eight digits, presented in two rows of four characters. The digits were selected as follows. First, the highest digit was selected randomly, ranging from 6 to 9. Then, the other seven distractors were selected randomly from the range 1 to the highest digit minus one. Tips meant that a target on one trial could be a distractor on another trial, and ensured no repetitions of the highest digit within a single display. The eight characters were assigned randomly to the fixed positions in the display. The mask display consisted of eight X's presented in the positions formerly occupied by the digits. The digits and masks measured about 0.3 cm width by 0.4 cm height, and the total display measured 3.9 x 3.0 cm, or 3.72 x 2.86 deg visual angle, based on a typical viewing distance of 60 cm. The characters were presented in white on a black background.

Design. The experiment was divided into 15 blocks of 30 trials each. Three different SOAs separated S1 (the tone) and S2 (the digits): 50, 150, and 650 ms. These were used equally often. Each block of 30 trials thus consisted of 10 trials at each SOA, presented in random order.

Procedure. The subjects were given written instructions describing the task. The instructions stated that the tone response should be made as rapidly and accurately as possible, while accuracy only was stressed on the second (highest digit) task. To discourage hasty responses to the digits, the program did not accept digit responses until 700 ms after the first response, requiring repetition of any prior to data collection, each subject worked through 72 practice trials, in 3 miniblocks of 24 trials each.

Figure 3 shows the procedure. Each trial began with the presentation of a plus sign as a fixation point, appearing at the center of the display for 1000 ms. Five hundred milliseconds after its offset, the tone (S1) was presented, at either 300 or 900 Hz for 150 ms. After the SOA of 50, 150, and 650 ms had elapsed, the display of digits appeared in the center of the screen, replaced with masks after the proper exposure duration. The subject responded to the tone by pressing either the Z or X key on the keyboard, corresponding to a low or a high tone, respectively, using the first or the second finger of their left hand. The subject responded to S2 (the digits) by pressing the B, N, M, or comma keys, corresponding to 6, 7, 8, or 9, respectively. As soon as the second response was detected by the computer, the display of masks was terminated. If an error was made on either task, a warning message ("ERROR!") was displayed for 750 ms, followed by a 250-ms offset. The intertrial interval between the second response and onset of the next fixation point was 1.3 s. At the end of each block the subject rested until ready to resume. Feedback was then provided, consisting of mean correct RT for the tone task, and number of errors on both the first and the second tasks.

The exposure duration for the digits was constant throughout a block, but adjusted between blocks. When accuracy on this task fell below 60% on a given block, the duration was

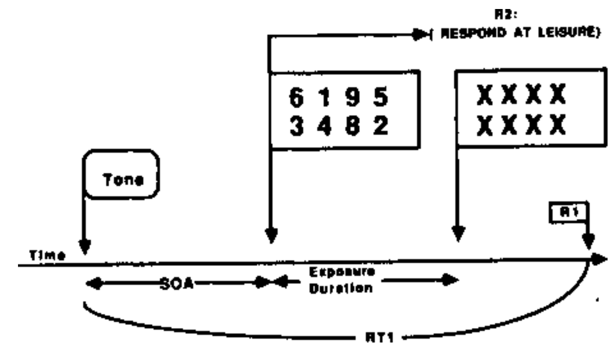

FIG. 3. The basic paradigm in Experiment 1. R1 is speeded, while R2 is made at the subjects' leisure; depending on the SOA, the tasks may overlap extensively.

increased by 17 ms on the following block, and if it exceeded 80%, the duration was decreased correspondingly. Thus, exposure duration was never confounded with the SOA between S1 (the tone) and S2 (the digits).

Results and Discussion

Basic Results

The data collection produced 4500 pairs of responses for each of the three SOA conditions (30 subjects x 150 response pairs). Mean R1 (tone response) RTs and percentage errors on the second task are presented in Fig. 4 as a function of the SOA. For this purpose, R1 times under 160 ms, or over 1000 ms, were discarded as deviant; the median number of discarded trials per subject was 7. Response times to the tone did not differ greatly as SOA was lengthened (475, 469, and 487 ms for SOAs of 50, 150, and 650 ms, respectively). An Anova showed that the effect of SOA was not quite significant, $F(2,58) = 2.5$, $.05 < p < .10$.

The percentage errors in the second (highest digit) task were 36.4, 36.2, and 33.9, for SOAs of 50, 150, and 650, respectively. This effect of SOA was significant, $F(2,58) = 3.5$, $p < .05$. The standard error on the observed difference is 2.1% (95% confidence level). The mean exposure duration for the digit display was 269 ms.

The mean percentage errors in the first (tone) task were 3.4, 2.9, and 3.3, for the SOAs of 50, 150, and 650, respectively. This effect was significant, $F(2,58) = 9.8$, $p < .001$.

Comments on the Basic Results

The results show that the decrease in SOA from 650 ms (minimal overlap) to 50 ms (maximal overlap) produces a statistically significant but minimal effect on accuracy in the second (digit) task. Several reasons for regarding this effect as minimal will emerge below (see especially Discussion of Experiments 1-4). The absence of substantial effects on R2 accu-

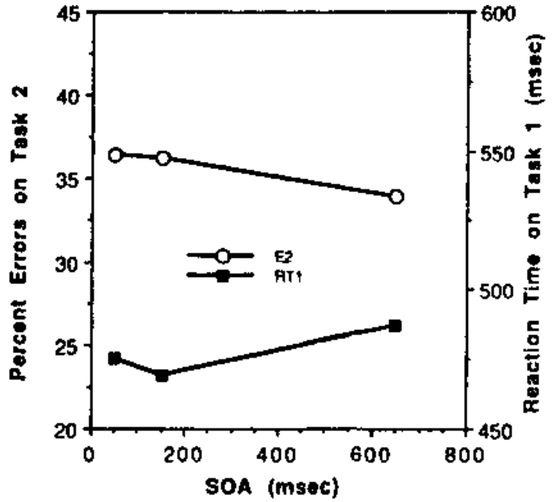

FIG. 4. Experiment 1: Error rates on task 2 (E2) and response times on task 1 (RT1), as a function of stimulus onset asynchrony (SOA).

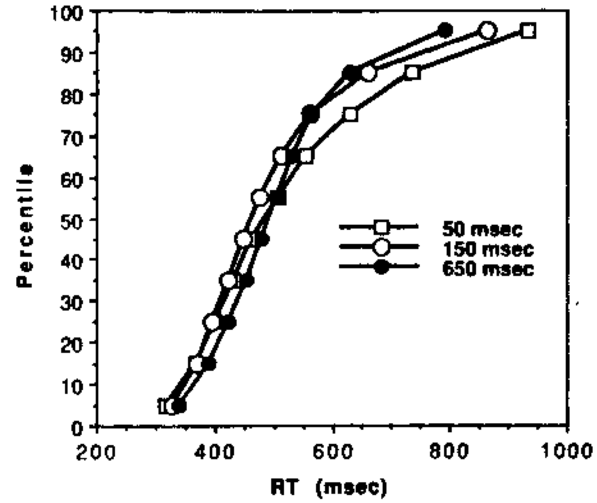

FIG. 5. Experiment 1: Cumulative latency distributions for R1 as a function of SOA.

racy confirms the predictions of the two-component theory, which claims that processing S2 through stimulus identification can proceed without interference under circumstances like these (in which the S1 does not also involve complex visual discriminations).

Reaction Time Distributions

To further investigate the effects upon R1 latencies (tone responses), the RT distributions were analyzed. First, for each subject, the approximately 150 correct response times from each SOA condition were rank ordered, and the scores approximating the 5th, 15th, . . . , 95th percentiles were estimated, using linear interpolation when necessary. No cutoffs were employed in this analysis. Then, these 30 percentile scores (3 SOAs x 10 percentile values) were averaged across subjects; i.e., the cumulative distribution functions were Vincentized. The results are graphed in Fig. 5 showing percentile as a function of RT for the three SOAs separately. The 50 and 150 SOA conditions differ little, and mostly in the slower responses. The 650 SOA condition produced a tighter distribution about its mean than the others. How can this be interpreted? The effect on the slower RTs may reflect occasional disruptions of fluent responding which occur more often at the shorter SOAs. The fact that the fastest RTs are actually *faster* at the short SOAs implies that the second stimulus may have a slight tendency to serve as a response accelerator, a phenomenon occasionally noted in the past (see Nickerson, 1967).

Dependencies between Tasks

The accuracy of second-task (highest digit) performance appears to have suffered only very slightly as the SOA is reduced. We can examine this further by looking at how accuracy in task 2 may vary depending upon the speed of the corresponding R1. If perceptual processing on the second task waited for completion of central processes in the first task, a very strong positive dependency should occur: fast R1s would be associated with accurate R2s, because on these trials the first task "got out of the way quickly." Furthermore, this should interact with SOA. By contrast, if both measures reflected the amount of capacity allocated to each task, with variability in the allocation ratio from trial to trial, then a *tradeoff* should appear: faster R1 times associated with less accurate R2 responses. However, if the two-component theory sketched above is correct, accuracy on task 2 and speed on task 1 are determined by entirely separate mechanisms—hence, no significant dependencies would be expected.

To examine these dependencies, one could simply compute correlations between performance in the two tasks, for each SOA. However, correlations are highly sensitive to extreme values, and could be misleading about any possible dependencies that might vary across the range of R1 latencies. Therefore, the accuracy of R2 conditional on the speed of R1 was computed. For each subject and SOA, the trials were ordered according to the R1 latencies. Then, for each quintile of this R1 distribution, the proportion of errors on the R2 responses that accompanied those

R1s was computed. Figure 6 shows the mean R2 error rates averaged across subjects for SOAs 50, 150, and 650, as a function of R1 quintile. The effect of R1 quintile was significant, $F(4,116) = 6.7$, $p < .001$, indicating a slight positive relationship between R1 latency and R2 error rates. More importantly, however, this trend does not differ according to SOA, $F(8,232) = 0.72$, $p > .60$. These effects can be accounted for on the plausible assumption that while the performance-determining aspects of both tasks operate independently, on some trials, the subject is simply poorly prepared for the entire task ensemble. For example, on some trials the subject is, to some degree, "out to lunch." If the weak dependency that is present were due to interference or waiting at the perceptual stage, it should grow as SOA was reduced, contrary to the results.

Summary

The results of this experiment provide support for response selection postponement: there is no sign that complex visual pattern processing on the second task suffers to any significant degree when it is forced to overlap with the first task. The two-component account actually predicts a dissociation: minimal effects on accuracy, but major effects on speed. Thus, a fuller discussion of these results awaits the next several experiments, which include some speeded second tasks.

EXPERIMENT 2

The results of Experiment 1 indicated that accuracy of responses to the second display were minimally affected by temporal overlap with the first task, requiring a speeded response. We now ask whether the same functional independence between a speeded first task, and perceptual processing on the second task, will arise when accurate feature conjunction is required. Treisman and Gelade (1980) argued that detection of conjunctions of features requires serial deployment of focal attention, based largely on the function relating RT to number of items in the display. The interpretation of these functions was questioned recently by Pashler (1987a), who observed that the slope pattern indicating serial self-terminating search arises only with very large displays that are usually searched with eye movements. Nonetheless, Treisman's interpretation seems to enjoy wide acceptance at the present time, and furthermore, the data undoubtedly *do* indicate profound performance limitations of a sort not found with search for single-feature targets (see Pashler (1987a) for an alternative account, however). Therefore, it is important to know the relationship between these visual performance limits, on the one hand, and the multitask attentional limitations, on the other hand. A possible account, entertained above, is that both might reflect the operation of the same mechanism—the "all purpose CPU."

The present experiment substitutes conjunction search (detecting a green T among green O's and red T's) for the highest-digit task of Experiment 1. Does temporal overlap with a speeded first task interfere with conjunction search accuracy? Is performance interdependent in the two tasks?

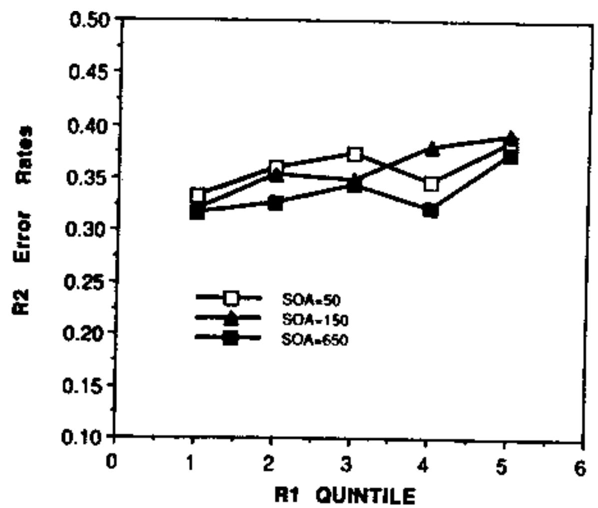

FIG. 6. Experiment 1: Mean R2 accuracy as a function of quintile where the R1 fell within R1 distribution (for that subject x SOA).

Method

Subjects. Thirty undergraduates at the University of California, San Diego, participated as subjects in the experiment, in partial fulfillment of a course requirement.

Apparatus and stimuli. The apparatus and stimuli were identical to those of Experiment 1, except for the nature of the second display. This display consisted of 8 green and red O's and T's, presented in the same size and in the same positions as the digit display of Experiment 1. The colors on the Sigma Designs monitors appeared highly saturated and discriminate. Displays were created by randomly filling each of the 8 positions with a green O or a red T (independently), and then, for positive displays, replacing a randomly selected item with a green T. The mask characters were X's, as in Experiment 1, but yellow.

Design. The design was like that of Experiment 1, except for the additional factor of the presence/absence of the conjunction target. Half the displays contained a target, and half did not.

Procedure. The procedure was basically the same as that of Experiment 1. The exposure durations used during practice, the initial exposure duration during the experiment proper, and the procedure for adjusting exposure durations between blocks were all identical to those employed in the earlier experiment. Subjects responded to target present and target absent displays by pressing the M and comma keys with the index and middle fingers of their right hand.

Results

The data collection produced 2250 pairs of responses for each of the six SOA conditions (30 subjects x 75 response pairs). Figure 7 presents subjects' mean RTs for correct responses to the tone and accuracy of responses in the conjunction search, as a function of the SOA between the tone and the digits. (Deviant RTs were discarded as in the previous experiment; the median number of trials removed per subject was 4.)

The response times to the tone were 438, 428, and 435 ms, for SOAs of 50, 150, and 650 ms, respectively. The effect of SOA was not significant, $F(2,58) = 2.7$, $.05 < p < .10$. The mean tone RTs were 433 and 434 ms when the second display did or did not contain a conjunction target, respectively; this effect was not significant, $F(1,29) = .05$, $p > .80$. The interaction of these was also nonsignificant, $F(2,58) = 1.9$, $p > .15$.

The mean percentage errors in the second (conjunction search) task were 23.9, 23.1, and 21.3, for SOAs of 50, 150, and 650, respectively. This effect of SOA was significant, $F(2,58) = 4.9$, $p < .02$. The difference of 2.6% between SOAs of 50 and 650 ms has a standard error of ± 2.2 % at the 95% confidence level. Overall accuracy did not differ as a function of whether a target was present or not (21.8% errors vs. 23.6%). There was a trend toward an interaction between the effects of SOA and target presence/absence, $F(2,58) = 3.0$, $.05 < p < .10$. This happened because at the short SOAs (50 and 150), responses were more accurate when a target was present (by 2.9 and 4.0%, respectively), but at the longer SOA,

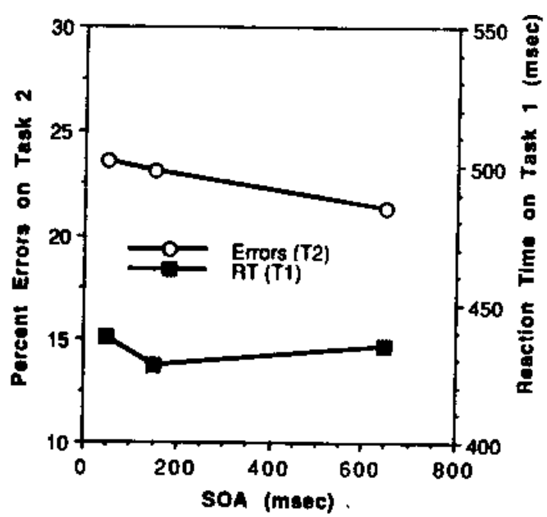

FIG. 7. Experiment 2: Errors on task 2 and RT on task 1, as a function of stimulus onset asynchrony (SOA).

responses were 1.6% more accurate when the target was *absent.* In short, the long SOA produced a reduction in the false alarm rate. The mean S2 exposure duration was 166 ms.

The mean percentage errors in the first (tone) task were 3.8, 3.3, and 2.3, for SOAs of 50, 150, and 650, respectively. This effect was significant, $F(2,58) = 8.1$, $p < .002$. The errors did not quite differ significantly as a function of presence vs. absence of a conjunction target (3.4 and 2.8%, respectively, $F(1,29) = 4.1$, $.05 < p < .10$.

Figure 8 shows the mean R2 error rate for SOAs of 50, 150, and 650, as a function of where the corresponding R1 lay within its distribution (by quintile). As in Experiment 1, there appears to be some very weak positive dependency (faster R1s associated with accurate R2s), but it is not significant, $F(4,116) = 2.1$, $.05 < p < .10$. As in the earlier experiment, this effect is not increased at the shorter SOAs, $F(8,238) = .96$, $p > .40$. Again, therefore, the pattern of results does not suggest any postponement of the perceptual processing of S2.

Discussion

The results are very similar to those of Experiment 1. Responses to the tone were barely affected by temporal overlap with the feature conjunction task, while accuracy in detecting conjunctions was impaired only very slightly as the SOA was shortened. Furthermore, performance on the two tasks was largely independent, and the weak positive contingency that was present was at least as large as the 650-ms SOA as it was at

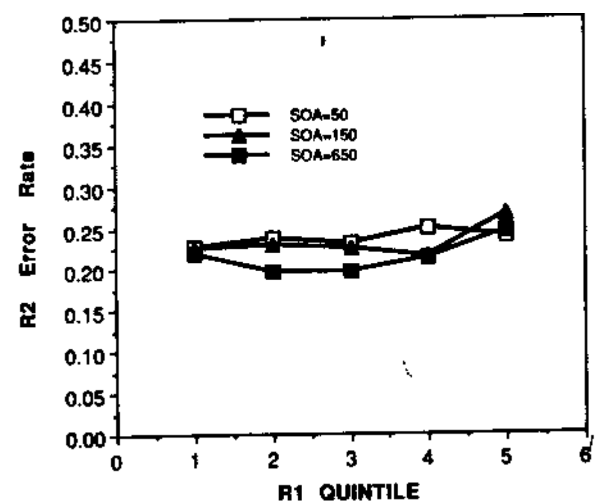

FIG. 8. Experiment 2: Mean R2 accuracy as a function of quintile where the R1 fell within R1 distribution.

shorter SOAs. Hence, the perceptual processing required for detection of feature conjunctions (which, according to Treisman and Gelade, involves serial scanning of focal attention) does not use the same mechanism and/ or capacity as performance of the first task, with its speeded response selection and execution.

Note that the first task latencies in this experiment are substantially faster than those in the previous experiment. The task is the same in both cases, and the second tasks (which do differ) do not appear to be having much impact on first task performance. How can this be? This probably illustrates the often overlooked role of preparation in determining (or, often, masquerading as) dual-task interference (on this, see Gottsdanker, 1980; Logan, 1978; Pashler, 1984). In the first experiment, subjects have to maintain a more complex second-task S-R mapping in mind, whereas in the second task, the mapping is much simpler. Pashler (1989) provides other illustrations of how overall performance (but not performance as a function of SOA) is affected by the entirety of the task specifications that must be maintained in any given dual-task condition.

The results of these experiments also have implications for the observations of Duncan (1980a). Duncan examined visual search performance with brief displays presented either simultaneously or successively, sometimes including two targets in displays. With simultaneous displays, subjects were more accurate to detect one target when they did not detect the other target. However, using small displays, when only a single target was presented, accuracy was the same with simultaneous and successive displays. Duncan's conclusion was that while the visual processing necessary to discriminate targets from distractors is parallel and capacity unlimited, when a target is detected, a limited capacity mechanism is required to enable a response. In Duncan's terminology, this mechanism is required whenever targets are to be transferred to the "second level." In the experiments reported here, second task performance must have required this process of taking note of a target, and thus entry into Duncan's second level (except perhaps for target-absent trials in Experiment 2). Yet performance was unaffected by the processes occurring on the first task (which were nonetheless sufficient to dramatically delay second-task responses, as shown by the results of Experiments 3 and 4). Thus, Duncan's limited capacity system must not be the same as the response selection bottleneck observed in PRP paradigms. What is it then? One reasonable possibility is that Duncan's phenomenon may arise only within a single modality (for the auditory equivalent, see Ostry, Moray, & Marks, 1976). Duncan's phenomenon may be due to narrowing of visual attention onto the target (observed by Hoffman, Nelson, & Houck, 1983). Cross-modal detection experiments would be necessary to resolve this question.

EXPERIMENT 3

In the preceding experiments, a rapid choice response to a tone was followed by an unspeeded judgment on a complex visual S2, which was masked at a constant interval from its onset. SOA effects upon accuracy in the second task were minimal, and additionally, there were no signs of meaningful dependencies between performance on the two tasks. The two-component theory predicts this, claiming that even demanding second-task perceptual processing need not wait for any stages in the speeded first task.

if this account is correct, massive effects on task 2 performance, as a function of SOA, should appear when R2 *latencies* are examined. With rapid selection and execution of R2 required, reductions in SOA should dramatically increase response times (the PRP effect). Furthermore, strong statistical dependencies should emerge, interacting with SOA. Experiment 3, therefore, employed stimuli and tasks identical to those of Experiment 1—the only difference was that the second display was not masked, and subjects were required to produce both responses as quickly as possible.

Method

Subjects. Fourteen undergraduates at the University of California, San Diego, participated as subjects in the experiment, in partial fulfillment of a course requirement.

Apparatus and stimuli. The apparatus and stimuli were identical to those of Experiment 1, except that there was no mask.

Design. The design was identical to that of Experiment 1.

Procedure. The procedure was identical to that of Experiment 1, except for the following changes. First, the subjects were given instructions in writing describing the task. In this experiment, these instructions stated that the first response should be made as quickly as possible, and that the second response should then also be made as quickly as possible. As in Experiment 1, each subject worked through 72 practice trials, in 3 miniblocks of 24 trials each. The second change pertained to the digits (S2): the display remained present on the screen until both responses had been made.

Results and Discussion

The data collection produced 1800 pairs of responses for each of the three SOA conditions (12 subjects x 150 response pairs). S1 (tone) response times under 160 ms, or in excess of 1000 ms, were discarded as deviant. Similarly, S2 (digit) response times under 200 ms, or in excess of 2200 ms, were discarded as deviant. The median number of trials removed per subject was 44, more than in the previous studies; note that this measure was conservative given the hypothesis being tested. Figure 9 presents subjects' mean RTs for correct responses to the tone and also to the digits, as a function of the SOA between the tone and the digits.

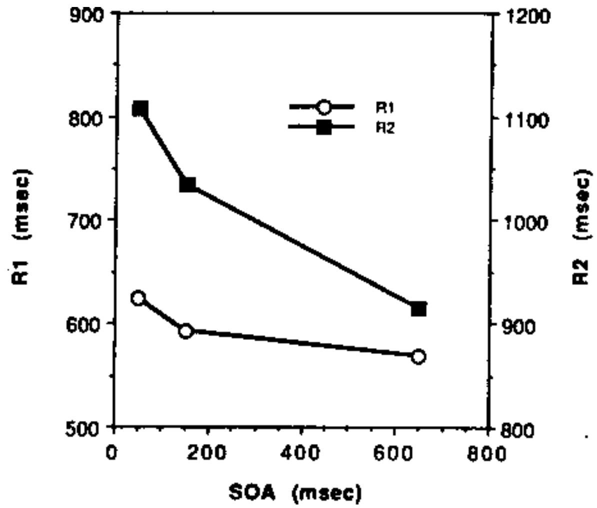

FIG. 9. Experiment 3: Mean RTs for R1 and R2 as a function of stimulus onset asynchrony (SOA).

The mean correct R1 latencies were 625, 594, and 568 ms, for SOAs of 50, 150, and 650 ms, respectively, a significant effect, $F(2,26) = 7.4$, $p < .01$.

The mean percentage errors in the first (tone) task were 1.9, 1.2, and 1.4, respectively. This effect was not significant, $F(2,26) = 2.7$, $p > .05$.

The mean correct R2 latencies were 1109, 1034, and 914 ms, for SOAs of 50, 150, and 650 ms, respectively. The effect was highly significant, $F(2,26) = 159.4$, $p < .001$.

The percentage errors in the second (highest digit) task were 8.8, 9.0, and 11.7, for SOAs of 50, 150, and 650, respectively. This effect of SOA was significant, $F(2,26) = 4.2$, $p < .05$.

Comments on the Basic RT Results

The results of the present experiment show a classic PRP function. Reducing the SOA increases R2 times, increasingly as SOA becomes shorter. Thus, reducing the SOA from 150 to 50 ms produces a 75-ms lengthening in the second response time—close to the classic "minus-one slope" predicted by the single-channel models (Welford, 1958). This also fits with the claim of the two-component model that performance of a critical stage in a speeded task 2 is subject to postponement on virtually all these trials. Recall, however, that in Experiment 1 the same reduction of SOA in the identical task situation produced very little change in the accuracy with which subjects could perform the same task, when the display was brief and masked. The results strongly support the view that interference between these two tasks has its locus beyond the stages of processing required for the identification of the digits.

Dependencies between Tasks in the RT Distributions?

The results just discussed indicate interference between the two tasks, as indicated by slowing of the response times when the two stimuli are brought closer together in time. Following the strategy in the previous two experiments, we can examine the dependencies between performance on the two tasks. Here we look at the way R2 speed, rather than R2 accuracy, depends upon the speed of R1. Figure 10 shows the mean R2 latency for SOAs 50, 150, and 650, as a function of which quintile the corresponding R1 latency fell within, among the R1 latencies for that subject and SOA. Unlike in the first two experiments, a dramatic dependency appears, specifically a tendency for R2 to be fast on trials where R1 was fast. The effect of R1 quintile on R2 latency was highly significant, $F(2,26) = 68.2$, $p < .001$. Furthermore, note that as SOA is reduced, the contingency becomes much more pronounced, $F(8,104) = 15.2$, $p < .001$. This is predicted by the theory diagrammed in Fig. 1, involving just response selection postponement, since reducing the SOA increases the proportion of the time on which the response selection in the second task waits for response selection on the first task to be completed. This model makes even more fine-grained predictions: at the longest SOAs, only the very slowest R1 times will produce slowing of R2, whereas at shorter

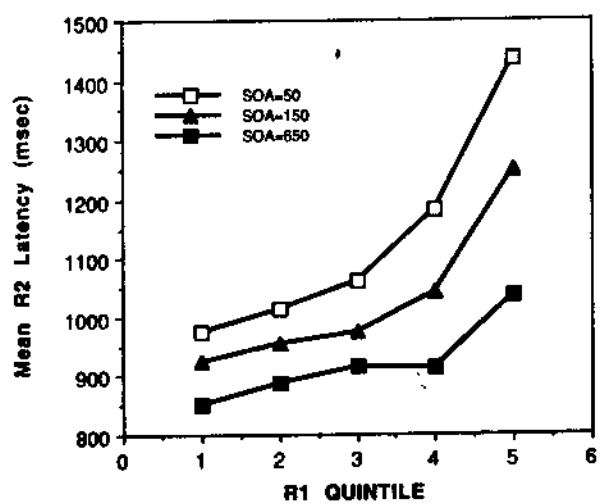

FIG. 10. Experiment 3: Mean R2 latency as a function of quintile where the R1 fell within R1 distribution.

SOAs, R2 elevation will appear farther down the R1 distribution. The data confirm this.

EXPERIMENT 4

In the preceding experiment, subjects made a rapid choice response to a tone, and then attempted to respond as rapidly as possible with the identity of the highest digit in an array of digits presented in varying temporal overlap to the performance of the tone task. The effects upon response times in the digit task were dramatic, whereas effects upon accuracy in the same task—with brief displays—were minimal in the first experiment. According to the theory sketched at the outset of the paper, the dramatic interference observed in Experiment 3 reflects response selection postponement, rather than a problem with programming and executing manual responses. Chronometric evidence for this was provided by Pashler (1984b) and Pashler and Johnston (1989).

If this account is correct, then the massive effect of SOA on task 2 speed should also appear if the subject is required to select and execute a *vocal* response in task 2. The magnitude of this interference should be quite similar to that observed in the previous experiment. However, there are some influential accounts, not previously discussed here, that suppose that this interference depends heavily upon *similarity of responses* (thus, see Allport, Antonis, & Reynolds, 1972; Allport, 1979; McLeod, 1977a, 1978; but see Pashler, 1989, for a rebuttal). These views would predict that changing the response modality should lessen or eliminate the SOA effects on R2 latencies. To test these alternative predictions, Experiment 4 used basically the same stimuli and tasks as those of Experiment 3—the only difference was that the subjects were required to produce a manual response to the tone, and a vocal second response of naming the highest digit.

Method

Subjects. Fourteen undergraduates at the University of California, San Diego, participated as subjects in the experiment, in partial fulfillment of a course requirement, although the data from only twelve were usable (see below).

Apparatus and stimuli. The apparatus and stimuli were the same as those in the previous experiment, except for the vocal responses. Subjects spoke through a DAK Industries ("Audio-telescope") highly directional microphone, which was plugged into a Gerbrandts Model G1341T voice-activated relay. The relay was in turn connected to the computer. The equipment was adjusted for each subject so that their vocal responses could be easily picked up, with minimal occurrence of spurious detections of keypresses.

Design. The design was identical to that of Experiment 3, except that there were 12 blocks of trials, instead of 15 blocks of trials (this to allow time for self-scoring of the vocal responses).

Procedure. The procedure was identical to that of Experiment 3, except for the following changes. When a vocal response was detected, the entire display disappeared, and was replaced with the correct vocal response (e.g., "eight"). The subject scored him or herself by pressing the space bar if the response was correct, and pressing the "/" key if the response was incorrect. In order to prevent hasty self-scoring, the machine did not accept scoring responses made before 650 ms had elapsed from the vocal response. Subjects were encouraged to be as accurate as possible. Previous experience indicated that subjects spontaneously detect virtually all of their errors in this task (see also Rabbitt & Rogers, 1977), and the present procedure seemed quite adequate to provide a reasonable estimate of the incidence of errors; the inevitable slight underestimation of errors could not affect the validity of our conclusions.

Results and Discussion

Despite the precautions mentioned above, for a few subjects microphone pickup of keypress responses was a problem. To detect these artifacts, we examined the incidence of vocal responses detected within 70 ms of the manual response. Given the nature of the tasks and RT distributions, such responses should properly occur only very rarely, and probably never for the longest SOA. As it happened, all the subjects showed at most a handful of such occurrences over the entire experiment, except for two subjects, who showed 28 and 148 of them. These two subjects were dropped from further analyses, and for the remaining subjects, the handful of trials fitting this artifact rejection criterion were excluded from the distributional analysis (but not from the computation of means). The data collection produced 1440 pairs of responses for each of the three SOA conditions (12 subjects x 120 response pairs). As in the previous experiment, S1 (tone) response times under 160 ms, or in excess of 1000 ms, were discarded as deviant. Similarly, S2 (digit) response times under 200 ms, or in excess of 2200 ms, were discarded as deviant. The median number of discarded trials per subject was 4.

Figure 11 presents subjects' mean RTs for correct responses to the tone and also to the digits, as a function of the SOA between the tone and the digits.

The mean correct R1 times were 491, 473, and 465 ms, for SOAs of 50, 150, and 650 ms, respectively, an effect that was significant, $F(2,22) = 8.9, p < .001$.

The mean percentage errors in the first (tone) task were 1.5, 1.5, and 1.4, respectively. This effect was not significant, $F(2,22) = 0.05, p > .90$.

The mean response times to the digits were 978, 904, and 768 ms, for SOAs of 50, 150, and 650 ms, respectively. This difference was highly significant, $F(2,22) = 49.9, p < .001$.

The percentage errors in the second (highest digit) task were 7.9, 8.4, and 9.4, for SOAs of 50, 150, and 650, respectively. This effect of SOA was significant, $F(2,22) = 7.7, p < .005$.

Comments on the Basic RT Results The results of the present experiment again show a classic PRP func-

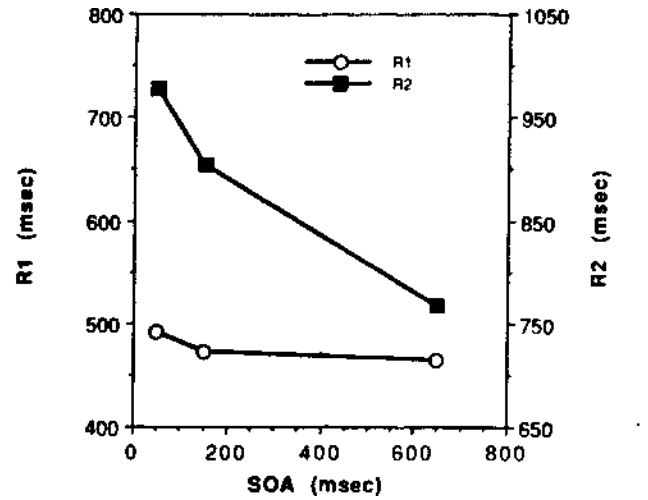

FIG. 11. Experiment 4: Mean RTs for R1 and R2 1, as a function of stimulus onset asynchrony (SOA).

tion. Reducing the SOA from 150 to 50 produced a 74-ms lengthening in the second response time, very close to the 75 ms observed in the previous study with two manual responses. The SOA reduction from 650 to 150 produced a 136-ms lengthening in RTs, close to the 117 ms of the previous experiment.

Dependencies between Tasks

The results just discussed show profound interference between the two responses, as indicated by the behavior of the R2 times when the two stimuli are brought closer together in time. Following the strategy in the previous experiment, the dependencies between performance on the two tasks were examined. Figure 12 shows the mean R2 latency, for SOAs of 50, 150, and 650, as a function of the quintile within which the corresponding R1 fell along the distribution of R1 times for that subject and SOA. The results are very similar to those of Experiment 3. A strong dependency is present, shown in the effect of R1 quintile on R2 latency, $F(4,44) = 32.4$, $p < .001$. Furthermore, note that the effect of R1 quintile grows as SOA is reduced: the interaction of SOA with R1 quintile is significant, $F(8,88) = 10.7$, $p < .001$. These observations closely fit predictions of the two-component theory outlined in the Introduction. Finally, note that at the longest SOA, it is just the very slowest R1 times that are associated with an elevation of R2, whereas at the shorter SOAs, the effect of R1 quintile on R2 times starts at lower and lower R1 quintiles. This pattern (present in this experiment, and also the preceding one) is

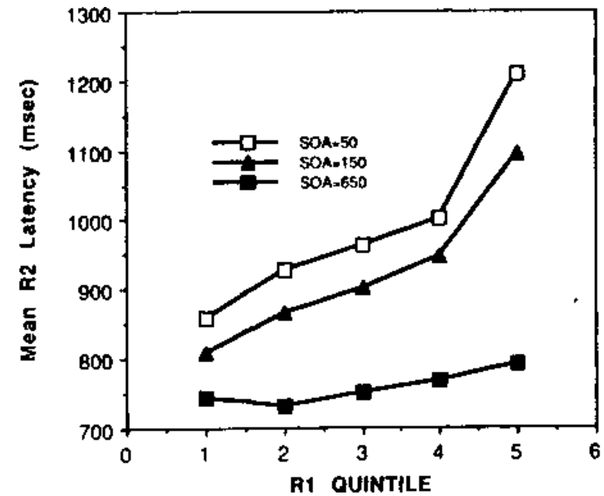

FIG. 12. Experiment 4: Mean R2 latency as a function of quintile where the R1 fell within R1 distribution.

precisely what the postponement model must predict, and it seems difficult to imagine how any competing model could make these detailed predictions.

DISCUSSION OF EXPERIMENTS 1-4

The results of the four experiments support the two-component theory sketched in the Introduction. When the first task involves a speeded choice response to a tone, and the second task involves a nonspeeded classification of a masked complex visual display, increasing temporal overlap of the tasks does not substantially impair accuracy of the second response. The result applies whether the second task requires determining the identity of the highest digit in the display (Experiment 1), or searching for a conjunction of color and form (Experiment 2). According to the two-component account (see Fig. 2), response selection in the first task postpones response selection in the second, but does not affect perceptual processing there. Thus, the likelihood of completing perceptual processing successfully before the mask terminates it is little affected by temporal proximity to the second task.

In discussing the results of Experiments 1 and 2, effects of SOA on R2 accuracy of approximately 2.5% were dismissed as minimal, despite their statistical significance. Is this justified? Clearly, the effects are small, but more importantly, their size and character show that they cannot be due to postponement of perceptual processing. First, the final SOA reduction from 150 to 50 did not seem to affect second-task accuracy at all, whereas in the speeded tasks, it produced the most direct effects on second-task performance. Second, if perceptual processing were delayed by anything like the 200-ms PRP effect, performance would plainly be drastically impaired, due to the masks. Could it be that these masks are somehow ineffective, and the perceptual processing is delayed but still does not suffer? To examine this unlikely possibility, the author recently required subjects to identify a single spatially probed item in masked displays virtually identical to these; delaying the probe by 200 ms produced accuracy impairments an order of magnitude larger than the SOA effects under discussion here. These conclusions receive further support from the examination of performance dependencies in Experiment 1 and 2. Dependencies are minimal, concentrated in the slower responses, and present at the long SOA as much as at the shorter ones.

By contrast, when the second task is speeded, and second task latencies are examined, we see dramatic delays that increase greatly as SOA is reduced; thus, the classic PRP effects of Experiments 3 and 4. Since response selection in the second task must wait for the completion of response selection in the first task, a strong positive dependency exists: faster R2s are associated with faster R1s. The probability of postponement on any given trial increases as the two tasks are squashed together in time, so the dependency increases dramatically as the SOA is reduced.

The two-component theory asserts that it is *selection,* not execution, of R2 that is postponed. Previous chronometric evidence supported this assertion: factors slowing second task response selection are additive with dual-task slowing and SOA (Pashler, 1984; Pashler & Johnston, 1989). Since these stages are subject to postponement, slowing them down delays the response as well. If it is indeed selection, rather than execution, of R2 that is postponed, then the delays should not depend upon the modality of response. The current results confirm this prediction: the SOA effects and intertask dependencies evident in Experiment 3 (manual R2) are very similar to those found in Experiment 4 (vocal R2). Note, however, that the manual responses (Exp. 3) are much slower in absolute terms than the vocal responses (Exp. 4). This is probably mostly due to the more difficult response selection required by the mapping from digits onto buttons in Experiment 3, compared to the highly prepracticed vocal naming in Experiment 4. The similarity of the SOA effects in these two experiments is basically another case of response selection delay factors being additive with SOA, and argues against the claim that single-channel effects depend on the use of the same response modality (Allport, 1979; McLeod 1977a, 1978). Pashler (1989) presents further evidence against this idea, examining manual versus vocal responses as a function of other factors.

In the Introduction, several competing accounts of the relation between

perceptual and multitask divided attention costs were mentioned. The present data not only support the two-component theory, they also raise serious problems for general capacity sharing models (e.g., Kahneman, 1973), and for the "CPU hypothesis" sketched in the Introduction. On the general capacity model, interference arises when one mental operation draws on a general pool of resources so heavily that the capacity available for others is depleted. This cannot explain the results of the first two experiments. The second tasks used there—highest digit naming and conjunction search—generate significant display size slopes in single-task visual search studies (Pashler & Badgio, 1985; Treisman & Gelade, 1980; see also Experiments 5 and 6 below). Treisman has argued that the conjunction search slopes reflect sequential scanning by a focal attention mechanism. If this mechanism is fueled by the putative general capacity, or actually carried out by some all purpose CPU, then accuracy should suffer dramatically when processing time is curtailed by the mask, and the SOA is reduced. No such impairments are observed.

Now, the pairs of tasks examined so far are not likely to produce perceptual interference, according to the two-component theory. We now turn to the situation in which both tasks require complex visual processing of multielement displays. Here, the theory predicts that a source of interference quite independent of response selection will arise, just as in the multistimulus experiments of Kleiss and Lane (1986). Reducing the SOA should now reduce R2 accuracy. In a sense, this prediction is of a successive advantage in a two-response version of the Shiffrin and Gardner paradigm. The two-component theory also makes further predictions that have not previously been examined. First, the reductions in R2 accuracy that do occur should not be related to R1 latencies, since according to the theory, the two components of interference are completely separate. Experiment 5 tests this prediction. Second, the SOA-induced accuracy decrement should not depend upon whether the first task requires speeded response selection and execution. As long as S1 is brief, its perceptual processing will commence immediately, thus impairing second-task accuracy, even when the subject is free to delay selection and execution of R1 (see Fig. 2). This is tested in Experiment 6, which differs from Experiment 5 in removing the requirement for a speeded first-task response.

EXPERIMENT 5

The purpose of the present experiment was to examine the effects of SOA on accuracy in the second (highest digit) task, as in the first experiment, but using a first task that requires complex visual processing. Here, S1 was a set of four diagonal line segments located to the left and right of the digit display: subjects had to decide if the slashes all pointed in one direction, or if instead there was one discrepant right-pointing slash. This task was chosen because it is perceptually demanding, but likely to require only featural discriminations, not processing of alphanumeric characters.

Method

Subjects. Twelve undergraduates at the University of California, San Diego, participated as subjects in the experiment, in partial fulfillment of a course requirement.

Apparatus and stimuli. The apparatus and stimuli were identical to those of Experiment 1, except for the nature of the first stimulus. Instead of a tone, a set of four slashes appeared on the outside corners of an imaginary square that was wider but shorted in height than the second display, for a duration of 150 ms. The outer dimensions of this display were 7.5 cm by 1.9 cm, or 7.13 by 1.81 deg visual angle, based on a typical viewing distance of 60 cm. Target-absent displays consisted of four left slashes (-45 deg clockwise from vertical) and target-present displays consisted of three left slashes and a single right slash (+ 45 deg). The position of the target slash in positive displays was chosen randomly. This second display and the mask for the second display were identical to those of Experiment 1.

Design. The design was identical to that of Experiment 1. The presence or absence of a target in the first display was determined randomly and independently on each trial.

Procedure. The procedure was identical to that of Experiment 1. The exposure durations used during practice, the initial exposure duration during the experiment proper, and the procedure for adjusting exposure durations between blocks were all identical to those employed in the earlier experiment. This served to equate conditions across experiments, but was not intended to (and did not) produce a comparable overall level of performance in this experiment.

Results

The data collection produced 1800 pairs of responses for each of the three SOA conditions (12 subjects x 150 response pairs). S1 response times under 160 ms, or in excess of 1000 ms, were discarded as deviant. The median number of trials per subject discarded was 8.

Figure 13 presents the response times to the first (visual) stimulus as a function of SOA, together with the accuracy of the highest digit response.

The first response times were 599, 572, and 588 ms, for SOAs of 50, 150, and 650, respectively. The effect of SOA was not significant, $F(2,22) = 2.5$, $p > .10$. The effect of target presence vs. absence (583 vs. 590, respectively) was also not significant, $F(1,11) = 0.2$, $p > .60$. However, the interaction of presence/absence with SOA was significant, $F(2,22) = 13.0$, $p < .001$. This reflects a reversal of the presence/absence effect; at the 50-ms SOA, yes responses are 37 ms slower, while they are 20 and 31 ms faster at the 150- and 650-ms SOAs, respectively.

The percentage errors on the first response were 11.6, 9.0, and 6.0, for SOAs of 50, 150, and 650, respectively. The effect of SOA was significant, $F(2,22) = 14.4$, $p < .001$. Neither the presence/absence effect nor the interaction of this with SOA was significant in this dependent measure.

The percentage errors on the second (highest digit) response was 61.0,

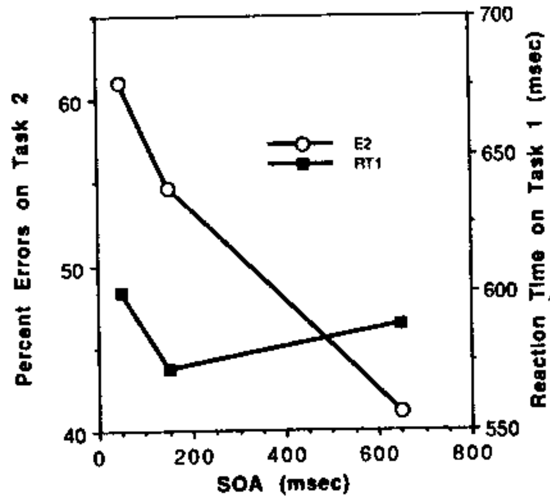

FIG. 13. Experiment 5: Errors on task 2 (E2) and RTs on task 1 (RT1), as a function of stimulus onset asynchrony (SOA).

54.6 and 41.1, for SOAs of 50, 150, and 650, respectively. The effect of SOA was highly significant, $F(2,22) = 60.5$ $p < .001$. The 19.9% difference as a function of SOA has a standard error of ± 5.4% at the 95% confidence level. The effect of target presence vs. absence in the first display was not significant, but the interaction of this factor with SOA was significant, $F(2,22) = 4.0$, $p < .05$. This appears to reflect the fact that accuracy was poorer when the first display had a target, but only for the SOAs of 50 and 650 ms; for the intermediate, 150-ms SOA, the effect was reversed. The mean S2 exposure duration was 300 ms.

Figure 14 presents mean R2 accuracy for each SOA, as a function of the quintile in which the corresponding R1 fell within its own distribution. The results show a weak effect of quintile, as in the first two experiments, $F(4,44) = 4.2$, $p < .01$. However, as in those experiments, quintile does not interact with SOA, $F(8,88) = 1.3$, $p > .25$. These results are again consistent with the view that the speed of R1 only affects R2 accuracy indirectly, because subjects are poorly prepared on some trials. A direct effect would predict the strong interactions between SOA and R1 quintile, as found in the dual latency analyses (Experiments 3 and 4). An additional analysis examined whether the proportion of *errors* on the first task differed as a function of whether or not the second response was correct. The percentage of errors on the first response was 10.6 given in accurate second response, and 9.2 given an error on the second response. This effect was not significant, $F(1,11) = 2.4$, $p > .15$, nor did it interact with SOA, $F(2,22) = 1.5$, $p > .20$.

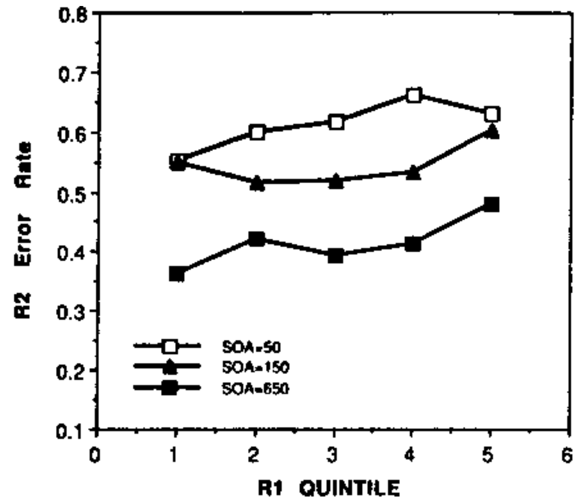

FIG. 14. Experiment 5: Mean R2 accuracy as a function of quintile where the R1 fell within the R1 distribution.

Discussion

This study yielded two primary results. The first is that when a complex visual first task is employed, a dramatic effect of SOA upon second-task accuracy arises. Experiment 1 involved a nonvisual first task and showed no such effect. Thus, the results of the first experiment cannot be dismissed by saying that performance on the accuracy task is "data-limited": limited capacity mechanisms are plainly required, but not ones that are "depleted" by overlapping the first task in the Experiment 1. The second result is that the SOA effects here are not accompanied by any strong dependencies between performance on the two tasks (as found in RTs when both tasks were speeded in Experiments 3 and 4). In short, temporal overlap hurts second-task accuracy quite substantially, but getting the first task out of the way quickly makes little difference. This absence of dependencies fits with the two-component account sketched at the beginning (see Fig. 2). In this conception, no common resource determines speed of R1 and accuracy of R2, nor is perceptual processing of S2 waiting on completion *of anything* in the first task. When each visual stimulus arrives, it grabs some portion of available visual processing capacity, and the completion of the first-task perceptual processing is essentially unrelated to speed on that task. Two underlying sources of interference are both at work, but they are fundamentally different: one is a bottleneck, but the other is not.

EXPERIMENT 6

The previous experiment showed that when the first task required a speeded response to a complex visual display, accuracy on the second (highest digit) task was impaired by temporal overlap. According to the two-component theory, this stems from just one of the two postulated causes of dual-task interference—the perceptual capacity limitation. If this interpretation is correct, then first-task response selection and execution are not responsible for second-task accuracy impairment. This suggests another nonobvious and testable prediction: if *both* responses are made at leisure, thus allowing postponement of response selection and execution in the first as well as the second task, impairment of second-task accuracy by SOA reduction should still be present, and with about the same magnitude.

Thus, Experiment 6 simply repeated Experiment 5, but removed the speed requirements on R1.

Method

Subjects. Twelve undergraduates at the University of California, San Diego, participated as subjects in the experiment, in partial fulfillment of a course requirement.

Apparatus and stimuli. The apparatus and stimuli were identical to those of Experiment 5.

Design. The design was identical to that of Experiment 5.

Procedure. The procedure was basically the same as that of Experiment 5, with a few exceptions. The instructions emphasized that speed was of no interest, and that the subjects should "take their time" and respond as accurately as possible on both tasks. In order to produce an analyzable number of errors on the first response, the duration of the first display was reduced to 50 ms.

Results

The data collection produced 1800 pairs of responses for each of the three SOA conditions (12 subjects x 150 response pairs). S1response times under 160 ms, or in excess of 1000 ms, were discarded as deviant. Figure 15 presents subjects' mean percentage of errors on the second task, as a function of the SOA between the first and second stimulus. The percentage of errors was 52.6, 44.6, and 34.1 , for SOAs of 50, 150, and 650, respectively, a highly significant effect, $F(2,22) = 40.8$, $p < .001$. The standard error on the 18.5% SOA effect is ± 4.2%, at the 95% confidence level. The mean exposure duration was 286 ms.

The percentage errors on the first task was 9.1, 4.8, and 2.7, for SOAs of 50, 150, and 650, respectively, which was also significant, $F(2,22) = 12.2$, $p < .001$.

Dependencies between Tasks

The accuracy of performance on the first task was assessed as a func-

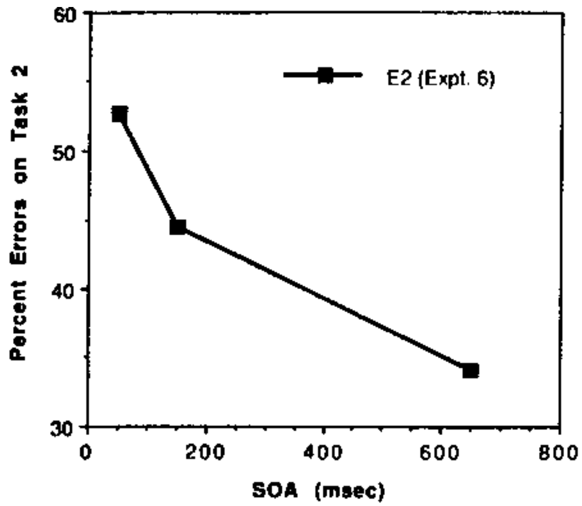

FIG. 15. Experiment 6: Errors on second task (E2) as a function of SOA.

tion of accuracy on the second task. Overall, the percentage of errors given a correct second response was 5.8, and given an error, 5.4. This difference was not significant, $F(1,11) = .18$, $p > .60$, nor did it interact with SOA, $F(2,22) = 1.6$, $p > .20$.

Discussion

The results show that accuracy on the second (highest digit) task is severely impaired as the SOA between that stimulus and the preceding visual stimulus is reduced, despite the fact that the subject could delay selection and execution of R1 in this experiment.

This effect appears very similar in magnitude to that observed in the preceding experiment, where speeded first responses were required. The results confirm that interference in the complex visual judgment is caused by the perceptual processing on the first task, not by the response selection and execution, since these were performed at leisure in this experiment.

Note that there is no dependency between accuracy achieved on the two tasks, even though the temporal proximity is harmful to both tasks. This confirms one of the key qualitative differences between the two components of divided attention costs postulated in the Introduction. When response selection queueing occurs, the relevant stage(s) of the second task are not delayed as much if the corresponding stage(s) of the first task can be completed quickly, hence the strong dependencies of Experiments 3 and 4. By contrast, the perceptual capacity limits exhibited

in Experiments 5 and 6 show no such dependencies, and accuracy in the second task is related to neither accuracy nor speed of the first response.

GENERAL DISCUSSION

The purpose of this article is to propose an account of how visual attentional limits are related to central attentional limits, and thus to provide the basic elements of a general theory of divided attention in simple tasks. According to the two-component theory proposed here, two separate and qualitatively quite different sorts of limitations constrain performance when people attempt to perform very simple tasks with visual stimuli presented at or near the same time. On the one hand, complex visual perceptual processes occur simultaneously, but they can generate mutual interference if their complexity reaches a certain level. On the other hand, central decision and response/selection operations cannot be performed simultaneously, and obligatory queueing occurs at this stage. This account differs greatly from previous (single- and multiple-) resource conceptions. First of all, these two limitations do not stem from any *single* pool of resources, and furthermore, neither the perceptual limit, nor especially the central limit, can be well characterized in terms of graded resource allocation.

The theory was broadly supported by evidence reviewed in the Introduction. Discrete central postponement was indicated by chronometric studies of the "PRP paradigm" (Pashler, 1984b; Pashler & Johnston, 1989), examining patterns of additive and underadditive interactions between dual-task slowing, SOA, and factors affecting different second-task stages. Evidence for visual capacity limitations was provided by studies of Prinzmetal and Banks (1983), Kleiss and Lane (1986), and others, using complex multielement arrays within a *single task.* Various details of these previous results provided hints that these two forms of limitation probably do not reflect aspects of a single common mechanism or capacity. Evidence suggested that visual capacity limitations do not reflect queueing or serial processing, and also that new stimuli can "grab" these resources without delay even if their order is unknown. By contrast, response selection postponement is inherently serial, and the delays it produces seem exacerbated when the order of stimuli is unknown (Pashler, 1989).

The two-component account generated many new predictions for dissociations and dependencies between speed and accuracy in two tasks performed at variable temporal asynchronies, with or without speed stress. The six experiments reported here appear to be the first to systematically examine dissociations between speed and accuracy in tasks requiring stimulus identification and choice responses, and the dependencies that arise between performance measures other than latencies (latency correlations in the PRP *have* been observed previously; see, e.g., Gottsdanker & Way, 1966; Pashler, 1984b).

The studies provided new and detailed support for the two-component theory, and, at least as importantly, they severely undermine competing models. First are dissociations in the effects of SOA on two different task 2 variables: response speed versus response accuracy (in unspeeded response to masked displays). When the first task is auditory, second-task accuracy is unaffected by SOA reduction, even when complex perceptual judgments are required on the second task. In Experiment 1, the judgment required the subject to determine the highest digit in a display of 8 digits, while Experiment 2 required detection of conjunctions of color and form. By contrast, when speeded second task responses were required to the highest digit task, latencies were dramatically increased when SOA was reduced. As the SOA was reduced from 150 to 50 ms, response delays were nearly as large as the SOA reduction, implying complete postponement over this range. Furthermore, delays were very similar whether the response was manual (Exp. 3) or vocal (Exp. 4).

The dissociation provides strong new evidence for the response selection postponement account (Pashler, 1984b; Pashler & Johnston, 1989; Welford, 1958). However, the last two experiments indicate that this is not to be explained by supposing that perceptual processing is "completely automatic," as late selection theorists have supposed, or "data-limited" rather than "resource-limited," as a proponent of general processing capacity might argue. The last two experiments show that when a complex visual discrimination is required on the first task, accuracy is reduced by approximately 20% as the SOA is shortened. This interference appears comparable whether or not the subject must make a speeded response on the first task, arguing further for the independence of central mechanisms (subject to queueing delays) and complexity-dependent visual interference. Furthermore, when the first response is speeded (Experiment 5), second-task accuracy is unrelated to the speed of the first response. This provides still further evidence that while there is a process in the first task which (largely) determines R1 latencies, for which second-task response selection must wait, perceptual processing on the second task does not have to wait for it at all.

Finally, the last two experiments offer support for the characterization of this perceptual interference proposed in the Introduction. Specifically, interference between complex perceptual processes may not be caused by visual stimuli competing to "grab" their share from a fixed pool of perceptual (or visual perceptual) "resources." If it were, the allocation would surely vary from trial to trial, and thus one would expect that accuracy on one task would tend to be associated with failure on the other

task. However, this sort of negative correlation was not observed in Experiment 6. This fits with the suggestion of Badgio and Pashler (1988) that "capacity sharing" at this stage may be a matter of mutual interference between simultaneous processes, generated perhaps by crosstalk, rather than true capacity sharing. Mozer (1987) describes results involving perception of multiple words that provide evidence for such underlying crosstalk, along with a connectionist model for the effect.

The fact that simultaneous complex visual processing with brief displays does not engender the performance tradeoffs expected on capacity sharing models converges with an interesting and neglected observation of Gardner (1973). Gardner required subjects to make two unspeeded two-alternative forced-choice judgments, one for each row of a brief masked display. Subjects were required to decide whether a T or an F was present in the top row, and whether a D or a reversed D was present in the bottom row. Each row contained zero or one confusable distractor element. When there was a confusable distractor on a given row, performance in the judgment relevant to that row was substantially impaired. However, performance on the *other row* was unimpaired. As Gardner pointed out, if presenting distractor elements increased the demands on a limited perceptual resource, *both* tasks should plainly be affected. Unfortunately, Gardner did not report the correlations between performance of one row and performance on the other, so his results might still be consistent with *a fixed* allocation of resources at the initiation of the trial.

Plainly, at the present time, we cannot claim to have fully characterized the perceptual processing limitation observed here. In particular, the suggestion made here that these limits arise when discriminations are "complex" or "difficult" is admittedly vague; further research will be needed to make this more precise. In any case, the goal of the present article has been to (1) argue that the perceptual processing limitation, and the response-selection limitation, are separate and fundamentally different in character, and (2) suggest that even the perceptual processing limitation may not really be much like a "resource" or "capacity."

The Multiple Resource Approach

The claim that divided attention costs do not reflect a single mechanism or capacity is not novel. Previous workers using the framework of capacity theory have stated as much (e.g., Gopher *et al*., 1982). Some have even proposed separate resources associated with different stages, with different kinds of processes, *and* with different response moralities (e.g., Wickens, 1983). If this is correct, it would not bode well for scientific progress. However, such conclusions have emerged from experiments that lack any real way to distinguish different possible mechanisms of interference in real time. Gross levels of performance achieved over several seconds or even minutes of performance of complex tasks have been recorded, as a function of task emphasis and overall task "difficulty." Latencies are collected in one task only, if at all, and the analyses have not addressed particular processing stages. Interference observed in such contexts could reflect switching, graded capacity allocation, buffering at input and output, response grouping, interference in short-term memory for intermediate computations, interference in task preparation (Gottsdanker, 1980), or any of the plethora of other phenomena. The capacity framework encourages amalgamating all these phenomena together with concepts like "capacity" or "resources." It is questionable whether any real empirical understanding can arise in this way. If the dependent measures are nondiagnostic, then applying even elegant formal analyses (e.g., from mathematical economics) cannot be expected to bring the underlying phenomena into clear focus. Not surprisingly, disenchantment seems to be arising of late among proponents of capacity theories (Navon, 1984). Carrying out more fine-grained, chronometric types of experiments, with tasks of manageable complexity, would seem to offer more promise for strong empirically based theories.

However, the present two-component account differs from multiple resource views not just in relying on a different sort of evidence, but also in reaching substantively different conclusions. The theory explicitly rejects genuine capacity sharing as a characterization of interference at response selection. The phrase "genuine capacity sharing" is meant to refer to any finite resource whose allocation is *graded* (a point about which Kahneman (1973) and McLeod (1977b) were explicit). The present work and the chronometric studies reported by Pashler (I984b) and Pashler and Johnston (1989) provide no support for a graded allocation of central resources; instead, they argue for a discrete queueing of response selection. Even the perceptual processing limitation, which has some resource-like properties, does not exhibit the microtradeoff that would be implied by the "capacity" metaphor. Why was genuine capacity sharing proposed in the first place? Mainly, it seems to have been the observation that, as subjects' priorities between tasks arc varied, one often sees a gradual tradeoff in performance: "graceful degradation" in performance of one task, as performance of the other improves. Pashler (I984b) proposed that this may often be a result of subjects *preparing* the underlying S-R mappings to different degrees (see also Logan, 1978). An apparent graceful tradeoff between tasks may in many cases simply reflect how much of the *time* immediately preceding the trial was spent rehearsing one mapping versus the other. If so, capacity models are highly misleading.

The Hoffman and Nelson Proposals

The present proposals were foreshadowed to some degree by Hoffman

and Nelson (1981). These investigators summarized several studies of visual attention tasks, with and without multiple response requirements. They proposed that visual attentional limitations were separate from response limitations, which were said to be stubborn and unaffected by visual variables (on this, see also Hoffman *et al.,* 1983). The present view disagrees with Hoffman and Nelson in attributing the response-related component to postponement of response *selection,* rather than response programming and execution. In addition, it characterizes the perceptual limitation rather differently. Nonetheless, the present work extends and validates the hunches those workers suggested.

Conclusions

The present article has defended a two-component theory of the limitations arising in simple divided attention tasks, postulating two components that are separate and quite different from each other (and also different in character from the limitations postulated by proponents of mental "resources"). The two-component theory can now claim detailed support from a range of converging empirical results. In Edition, these results would seem to rule out the obvious alternatives. Nonetheless, it is certainly possible that further research will require minor, or even major revisions of this conception. The author has little doubt that more complex and realistic dual-task situations will involve further sources of interference, in addition to those postulated here. Whatever degree of modification may be required, the findings indicate that when divided attention tasks are studied in fine-grained detail, using a wide range of available manipulations and response measures (particularly those derived from the study of processing stages; e.g., Sternberg, 1969; Salt house, 1981), rich empirical constraints on theorizing can emerge.

REFERENCES

Allport, D. A. (1979). Attention and performance. In G. Claxton (Ed.), *New directions in cognitive psychology.* London: Routledge and Kegan Paul.

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

Atkinson, R., Holmgren. J., & Juola, J. (1972). Processing time as influenced by the number of elements in a visual display. *Perception & Psychophysics,* **6**, 321-326.

Badgio, P. C., & Pashler, H. E. (1988). *What is the nature of capacity limitations in visual object recognition?* Manuscript submitted for publication.

Bertelson, P. (1966). Central intermittency twenty years later. *Quarterly Journal of Experimental Psychology,* **18**, 153-163.

Blake, R., & Fox, R. (1969). Visual form recognition threshhold and the psychological refractory period. *Perception & Psychophysics,* **5**, 46-48.

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

Broadbent, D. E. (1958). *Perception and communication.* London: Pergamon.
Broadbent, D. E. (1982). Task combination and the selective intake of information. *Acta Psychologica,* **50**, 253-290.
Duncan, J. (1980a). The locus of interference in the perception of simultaneous stimuli. *Psychological Review,* **87**, 272-300.
Duncan, J. (1980b). The demonstration of capacity limits. *Cognitive Psychology,* **12**, 75-96.
Duncan, J. (1983). Category effects in visual search: A failure to replicate the "oh-zero" phenomenon. *Perception & Psychophysics,* **34**, 221-232.
Eriksen, C. W., & Hoffman, J. (1973). The extent of processing of noise elements during selective encoding from visual displays. *Perception & Psychophysics,* **14**, 155-160.
Eriksen, C. W., & Spencer, T. (1969). Rate of information processing in visual perception: Some results and methodological considerations. *Journal of Experimental Psychology Monograph,* **79** (2, Part 2).
Gagnon, M., Cavanagh, P., & Laurencelle, L. (1978). Processing differences between memory search and foveal visual search. *Perception & Psychophysics,* **23**, 258-260.
Gopher, D., Brickner, M., & Navon, D. (1982). Different difficulty manipulations interact differently with task emphasis: Evidence for multiple resources. *Journal of Experimental Psychology: Human Perception and Performance,* **8**, 146-157.
Gordon, I. E. (1968). Interactions between items in visual search. *Journal of Experimental Psychology,* **76**, 348-355.
Gottsdanker, R. (1980). The ubiquitous role of preparation. In G. E. Stelmach & J. Requin (Eds.), *Tutorials in motor behavior.* Amsterdam: North-Holland.
Gottsdanker, R., & Way, T. C. (1966). Varied and constant inter-signal intervals in psychological refractoriness. *Journal of Experimental Psychology,* **72**, 792-804.
Green, B. F., & Anderson, L. K. (1956). Color coding in a visual search task. *Journal of Experimental Psychology,* **51**, 19-24.
Herman, L. M., & Kantowitz, B. H. (1970). The psychological refractory period effect: Only half the double-simulation story? *Psychological Bulletin,* **73**, 74-86.
Hoffman, J. E., Nelson, B., & Houck, M. R. (1983). The role of attentional resources in automatic detection. *Cognitive Psychology,* **51**, 379-410.
Kahneman, D. (1973). *Attention and effort.* New York: Prentice-Hall.
Kantowitz, B. (1974). Double stimulation. In B. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition* (pp. 83-131). Hillsdale, NJ: Erlbaum.
Keele, S. W. (1973). *Attention and human performance.* Pacific Palisades, CA: Goodyear.
Kleiss, J. A., & Lane, D. M. (1986). Locus of persistence of capacity limitations in visual information processing. *Journal of Experimental Psychology: Human Perception and Performance,* **12**, 200-210.
Kristofferson, L. (1972). Effects of practice on character classification performance. *Canadian Journal of Psychology,* **26**, 54-40.
Logan, G. (1978). Attention in character classification tasks: Evidence for the automaticity of component states. *Journal of Experimental Psychology: General,* **107**, 32-63.
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 & Performance,* **12**, 549-563.
McClelland, J. L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review,* **86**, 287-330.
McLeod, P. (1977a). A dual task response modality effect: Support for multiprocessor models of attention. *Quarterly Journal of Experimental Psychology,* **29**, 651-667.
McLeod, P. (1977b). Parallel processing and the psychological refractory period. *Acta Psychologica,* **41**, 381-391.

McLeod, P. (1978). Does probe RT measure central processing demand? *Quarterly Journal of Experimental Psychology,* **30**, 83-89.
Meyer, D. E., Yantis, S., Osman, A., & Smith, J. E. K. (1984). Discrete versus continuous models of response preparation: A reaction time analysis. In S. Kornblum & J. Requin (Eds.), *Preparatory states and processes.* Hillsdale, NJ: Erlbaum.
Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology,* **14**, 247-279.
Miller, J. (1988). Discrete and continuous models of human information processing: Theoretical distinctions and empirical results. *Acta Psychologica,* **67**, 191-257.
Mozer, M. (1987). Early parallel processing in reading: A connectionist approach. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading.* Hillsdale, NJ: Erlbaum.
Navon, D. (1984). Resources- a theoretical soup stone? *Psychological Review,* **91**, 216-234.
Neisser, U., Novick, R., & Lazar, R. (1963). Searching for ten targets simultaneously. *Perceptual and Motor Skills,* **17**, 955-961.
Nickerson, R. S. (1967). Psychological refractory phase and the functional significance of signals. *Journal of Experimental Psychology,* **73**, 303-312.
Norman, D. A., & Bobrow, D. J. (1975). On data-limited and resource-limited processes. *Cognitive Psychology,* **7**, 44-64.
Norman, D. A., & Shallice, T. (1985). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz, A D. Shapiro (Eds.), *Consciousness and self-regulation* (Vol 4). New York: Plenum.
Ostry, D., Moray, N., & Marks, G. (1976). Attention, practice and semantic targets. *Journal of Experimental Psychology: Human Perception and Performance,* **2**, 326-336.
Pashler, H. (1984a). Evidence against late selection: Stimulus quality effects in previewed displays. *Journal of Experimental Psychology: Human Perception and Performance,* **10**, 429-448.
Pashler, H. (1984b). Processing stages in overlapping tasks: Evidence for a central bottleneck. *Journal of Experimental Psychology: Human Perception and Performance,* **10**, 358-377.
Pashler, H. (1987a). Detecting conjunctions of color and form: Reassessing the serial search hypothesis. *Perception & Psychophysics,* **41**, 191-201.
Pashler, H. (1987b). Target-distractor discriminability in visual search. *Perception & Psychophysics,* **41**, 285-292.
Pashler, H. E. (1989). Do response modality effects support multiprocessor models of divided attention? *Journal of Experimental Psychology: Human Perception and Performance,* in press.
Pashler, H., & Badgio, P. (1985). Visual attention and stimulus identification. *Journal of Experimental Psychology: Human Perception and Performance,* **11**, 105-121.
Pashler, H., & Badgio, P. (1987). Attentional issues in the identification of alphanumeric characters. *Attention A performance XII: The psychology of reading.* Hillsdale, NJ: Erlbaum.
Pashler, H., & Johnston, J. C. (1989). Chronometric evidence for central postponement in temporally overlapping tasks. *Quarterly Journal of Experimental Psychology,* **41A**, 19-45.
Posner, M. (1982). Cumulative development of attention theory. *American Psychologist,* **37**, 168-179.
Prinzmetal, W., & Banks, W. P. (1983). Perceptual capacity limits in visual search. *Bulletin of the Psychonomic Society,* **21**, 263-266.
Rabbitt, P., & Rogers, B. (1977). What does a man do after he makes an error? An analysis of response programming. *Quarterly Journal of Experimental Psychology,* **29**, 727-743.

Rumelhart, D. E. (1970). A multicomponent theory of the perception of briefly exposed visual displays. *Journal of Mathematical Psychology,* **7**, 191-216.
Santee, J. L., & Egeth, H. E. (1982). Do reaction time and accuracy measure the same aspects of letter recognition? *Journal of Experimental Psychology: Human Perception and Performance,* **8**, 489-501.
Schweickert, R. (1978). A critical path generalization of the additive factor method: Analysis of a Stroop task. *Journal of Mathematical Psychology,* **18**, 105-139.
Shiffrin, R. M., & Gardner, G. T. (1972). Visual processing capacity and attentional control. *Journal of Experimental Psychology,* **93**, 72-83.
Smith, M. C. (1967). Theories of the psychological refractory period. *Psychological Bulletin,* **67**, 202-213.
Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. In W. G. Roster (Ed.*), Attention and Performance II* (pp. 276-315). Amsterdam: North Holland.
Townsend, J. T., & Ashby, G. (1983). *Stochastic modeling of elementary psychological processes.* Cambridge: Cambridge Univ. Press.
Treisman, A., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology,* **12**, 97-136.
van der Heijden, A. H. C. (1975). Some evidence for a limited-capacity parallel self-terminating process in simple visual search tasks. *Acta Psychologica,* **39**, 21-41.
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. (1980). The single-channel hypothesis. In A. T. Welford (Ed.), *Reaction time* (pp. 215-252). New York: Academic Press.
Wickens, C. D. (1980). The structure of attentional resources. In R. Nickerson (Ed.), *Attention and performance VIII.* Englewood Cliffs, NJ: Erlbaum.
Wickens, C. D. (1983). Processing resources in attention, dual task performance, and workload assessment. In R. Parasuraman & R. Davies (Eds.), *Varieties of attention.* New York: Academic Press.

(Accepted March 23, 1989)

COGNITIVE PSYCHOLOGY 21, 515-550 (1989)

Suppressing Natural Heuristics by Formal Instruction: The Case of the Conjunction Fallacy

FRANCA AGNOLI

University of Padova

AND

DAVID H. KRANTZ

Columbia University

A basic principle of probability is the conjunction rule, $p(B) \geq p(A\&B)$. People violate this rule often, particularly when judgments of probability are based on intensional heuristics such as representativeness and availability. Though other probabilistic rules are obeyed with increasing frequency as people's levels of mathematical talent and training increase, the conjunction rule generally does not show such a correlation. We argue that this recalcitrance is not due to inescapable "natural assessments"; rather, it stems from the absence of generally useful problem-solving designs that bring extensional principles to bear on this class of problem. We predict that when helpful extensional strategies are made available, they should compete well with intensional heuristics. Two experiments were conducted, using as subjects adult women with little mathematical background. In Experiment 1, brief training on concepts of algebra of sets, with examples of their use in solving problems, reduced conjunction-rule violations substantially, compared with a control group. Evidence from similarity judgments suggested that use of the representativeness heuristic was, reduced by the training. Experiment 2 confirmed these training effects and also tested the hypothesis that conjunction-rule violations are due to misunderstanding of "B" as "B and not A." Changes in detailed wording of the propositions to be ranked produced substantial effects on judgment, but the pattern of these effects supported the hypothesis that, for the type of problem used here, most conjunction errors are due to use of representativeness or availability. We conclude that such intensional heuristics can be suppressed when alternative strategies are taught.

We thank AT&T Bell Laboratories for supporting this research. We thank the following individuals for discussions and comments: MaryJo Altom, Harry Gollob, Jan Keenan, Kevin Miller, Nancy Pennington, Peter Poison, George Potts, Phil Shaver, Linda Roberts, and Ted Wright. We are grateful to Don Dulany, Doug Medin, and two anonymous reviewers for their insightful criticisms of earlier drafts, and especially to Buz Hunt, Julian Hochberg, and Steve Poltrock for extremely detailed and valuable comments and suggestions. The data reported were part of a doctoral dissertation submitted by the first author in partial fulfillment of the requirements for a Ph.D. degree at the University of Denver. Portions of the data were first reported at the Meetings of the Psychonomic Society in Boston; November, 1985. Correspondence and reprint requests should be addressed to Franca Agnoli, University of Padova, Via Beato Pellegrino 26, 35137 Padova, Italy.

0010-0285/89 $7.50