

The source of dual-task limitations: Serial or parallel processing of multiple response selections?

Suk Won Han · René Marois

Published online: 18 July 2013
© Psychonomic Society, Inc. 2013

Abstract Although it is generally recognized that the concurrent performance of two tasks incurs costs, the sources of these dual-task costs remain controversial. The serial bottleneck model suggests that serial postponement of task performance in dual-task conditions results from a central stage of response selection that can only process one task at a time. Cognitive-control models, by contrast, propose that multiple response selections can proceed in parallel, but that serial processing of task performance is predominantly adopted because its processing efficiency is higher than that of parallel processing. In the present study, we empirically tested this proposition by examining whether parallel processing would occur when it was more efficient and financially rewarded. The results indicated that even when parallel processing was more efficient and was incentivized by financial reward, participants still failed to process tasks in parallel. We conclude that central information processing is limited by a serial bottleneck.

Keywords Dual-task performance · Attention · Psychological refractory period (PRP)

When we attempt to carry out two arbitrary sensory–motor tasks in rapid succession, the response to the second task is almost invariably delayed (Welford, 1952), a phenomenon known as the psychological refractory period (PRP). Why do we show such severe limitations in performing dual tasks?

According to the serial bottleneck model (Pashler, 1984, 1994a, 1994b), this dual-task cost originates from the presence of a central bottleneck of information processing that allows only a single response selection operation to proceed at a time, thereby inducing serial postponement of response selection for the second task. By contrast, a broad range of “cognitive-control” models propose that multiple response selections can proceed in parallel (Meyer & Kieras, 1997; Miller, Ulich, & Rolke, 2009; Tombu & Jolicoeur, 2003), and that serial postponement of task performance is simply the result of individuals adopting a serial processing strategy.

Which classes of model best capture the source(s) of dual-task limitations? The cognitive control models make a simple prediction that would disconfirm the serial bottleneck model: If multitask performance is under control, it should be possible to execute two tasks in parallel. However, participants in dual-task conditions overwhelmingly exhibit a serial mode of information processing (Tombu & Jolicoeur, 2003). Some cognitive control proponents have attributed this finding to insufficient practice with the task conditions for developing the appropriate cognitive control necessary to perform tasks in parallel (e.g., Schumacher et al., 2001). However, whereas extensive practice may virtually eliminate dual-task costs (Schumacher et al., 2001; but see Tombu & Jolicoeur, 2004), it does not logically follow that this training effect results from the development of parallel processing of multiple response selections (Ruthruff, Johnston, & Remington, 2009). Practice could just as well lead to faster processing speed of each task, which could dramatically reduce the time that the second of the two tasks is postponed until it gets access to the bottleneck stage of response selection. Indeed, although recent behavioral, neuroimaging, and modeling work is consistent with the latter possibility (e.g., Dux et al., 2009; Kamienkowski, Pashler, Dehaene, & Sigman, 2011; Ruthruff, Johnston, & Van Selst, 2001), no direct evidence for the former has been found.

S. W. Han · R. Marois
Department of Psychology, Vanderbilt Vision Research Center,
and Center for Integrative and Cognitive Neurosciences,
Vanderbilt University, Nashville, TN, USA

S. W. Han (✉)
Department of Psychology, 627 Wilson Hall 111 21st Avenue
South Nashville, Tennessee 37240, USA
e-mail: suk.w.han@vanderbilt.edu

Although several previous studies have claimed that extensive practice should not only facilitate processing of each task, but also allow one to perform two tasks in parallel (Maquestiaux, Laguë-Beauvais, Bherer, & Ruthruff, 2008; Ruthruff, Hazeltine, & Remington, 2006), these studies cannot address whether parallel processing is possible at the response selection stage because the practice effects were claimed to result from bypassing the capacity-limited response selection stage (i.e., task automatization; see Maquestiaux et al., 2008) rather than from parallel processing at that particular stage. One can only test the theory of parallel processing at the response selection stage if both tasks engage that stage.

Another class of cognitive-control models—the *graded-sharing* models—also suggest that multiple response selections can proceed in parallel (Miller et al., 2009; Navon & Miller, 2002; Tombu & Jolicœur, 2003), but they differ from the parallel processing models described above in two important respects. Specifically, they argue that parallel processing can take place even in the absence of extensive practice and, second, that dual-task costs are inevitable (Tombu & Jolicœur, 2004). Although the second point appears consistent with the serial bottleneck model, the graded-sharing models argue that dual-task limitations do not arise from a serial bottleneck but rather from a capacity-limited central resource. In that framework, resources can be flexibly allocated to each task—which enables multiple response selections to proceed in parallel—but because these resources are limited the processing rate of each task is dependent on the amount of resources allocated to them (Navon & Miller, 2002; Tombu & Jolicœur, 2003). Viewed in this framework, serial processing is simply a special case of graded sharing, in which the proportion of capacity allocated to the first task (sharing proportion, or SP) is 100%.

Although they are powerful in their explanatory accounts, graded-sharing models beg the following question: If central resources can be flexibly allocated to each task, why, then, is serial processing so predominantly observed in dual-task situations? Proponents of parallel models have argued that the bulk of PRP studies included task instructions and/or contexts that biased the participants to adopt a serial processing strategy (Navon & Miller, 2002; Tombu & Jolicœur, 2003). However, serial postponement of task performance is still observed even when tasks are presented in a randomized order and equally emphasized (Dux et al., 2009; Pashler, 1994b), even when participants are rewarded to process both tasks in parallel (Ruthruff et al., 2009).

Although recent studies cast doubt on the task instructions/settings argument to explain the predominance of serial processing in dual-task situations, another argument has been more enduring: the general inefficiency of parallel as compared to serial processing. If one defines task efficiency in terms of the sum of RT to the two tasks, then serial processing

is more efficient and therefore favored over parallel processing. Even though parallel models differ in suggesting how capacities should be divided among tasks and how multiple response selections should proceed in parallel, it is generally agreed that such processing is less efficient than serial processing (Meyer & Kieras, 1997; Miller et al., 2009; Tombu & Jolicœur, 2003). In particular, according to the graded-sharing model, which can account for much of the extant PRP data (Fig. 1), parallel processing and serial processing predict similar Task 2 RT slowing with Task 1–Task 2 SOA. However, parallel processing also predicts slowing of Task 1 RTs—because capacity-limited resources are shared between the tasks—whereas serial processing predicts no effect of SOA on Task 1 RTs. This is because the graded-sharing model posits that amodal, central processing resources for response selection are flexibly allocated to meet task demands, such that processing resources allocated to the first task can be instantaneously reallocated to the second task when the first one is completed. This assumption enables the model to explain why Task 2 RTs should still depend on the Task 1–Task 2 RT SOA, even with parallel processing, which makes the model’s explanatory power excel that of other parallel models, as well as of the serial bottleneck model. By contrast, other parallel models predict that task RTs should be invariant across SOAs unless those SOAs are long enough to enable each task to be performed separately (e.g., Miller et al., 2009).

The task efficiency argument makes a very clear prediction: If participants adopt serial processing because it is usually the most efficient mode of performing two tasks in multitask conditions, then they should adopt a parallel

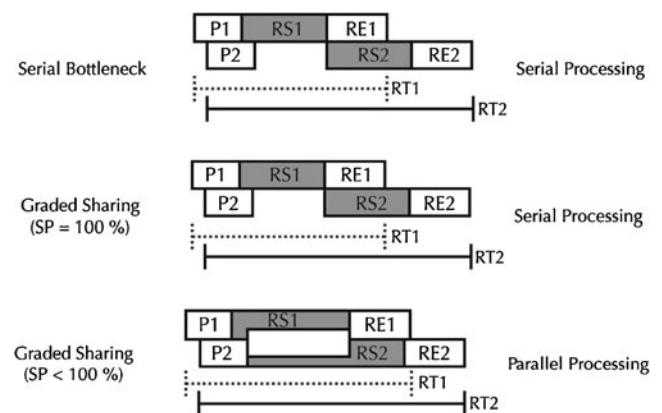


Fig. 1 Diagram depicting the graded-sharing and serial bottleneck models under short SOA conditions. With serial processing, only the Task 2 (T2) response is slowed, whereas in parallel processing Task 1 (T1) is also slowed, except when all processing resources are allocated to T1 (SP = 100%). Importantly, T2 is slowed to the same extent under both serial and parallel conditions. This is because the graded-sharing model posits that resources allocated to T1 become instantaneously available for T2 as soon as T1 is completed. Both models also posit that the perceptual identification (P) and response execution (RE) stages can proceed in parallel. RS, response selection; SP, sharing proportion; RT1, the first-task RT; RT2, the second-task RT

processing mode if this becomes the most efficient mode. If dual-task performance is instead limited by a serial bottleneck, serial processing will be maintained even when it is counterproductive, as the bottleneck allows only one response selection to proceed at a time. The purpose of the present study was to distinguish between these two possibilities. To do so, we created an experimental condition in which serial processing was rendered less efficient than parallel processing in terms of total processing time, and further incentivized participants to adopt parallel processing with financial reward. If processing resources can be shared between tasks, then participants should adopt a parallel processing strategy rather than a serial strategy because the former is now more efficient and financially rewarding. By contrast, the serial bottleneck model predicts that serial processing will still be maintained even when it is the counterproductive strategy.

To test these predictions, we used a dual-task experimental design inspired from Tombu and Jolicoeur (2003), which was conceived to obtain evidence for parallel processing. Specifically, an easy sensory–motor task was paired with a more demanding sensory–motor task, but with the easy task presented 300 ms after the demanding task. By instructing participants to respond to the stimulus of the easy task first, serial processing was rendered less efficient than parallel processing because only serial processing produces cognitive slack (Fig. 2). We also further incentivized participants to adopt parallel processing by providing financial reward.

The easy task was a two-alternative discrimination (2AD) visuo-vocal (VV) task, and the more demanding one consisted of an eight-alternative discrimination (8AD) audio-manual (AM) task. Both tasks were presented in each trial, with one of two possible SOAs (0 and 300 ms), which were either blocked (Exps. 1a and 2) or intermixed within a block (Exp. 1b). At the 0-ms SOA the visual and auditory stimuli were presented at the same time, whereas in the 300-ms SOA the auditory stimulus was always presented first.

The parallel and serial processing strategies make specific predictions about the task response times under the different SOA conditions (Fig. 2). These predictions assume that the central stages of the two tasks completely overlap with each other. That is, central processing resources are shared with the AM task during the entire response selection stage of the VV task. Under this assumption, the RTs for the VV task should be the same at both SOAs regardless of processing strategy. By contrast, the AM RTs at the 300-ms SOA should differ, depending on whether a parallel or serial processing strategy was adopted. Specifically, whereas the AM RT at the 300-ms SOA should be the same as that at the 0-ms SOA under parallel processing, serial processing should yield significantly extended AM RT at the 300-ms SOA, because

of the slack caused by the requirement to respond to the VV task first (Fig. 2). Hence, the comparison of the AM RT between SOAs would reveal whether serial or parallel processing had occurred.

Note that the VV RTs should not be affected by the different SOAs, though they should be longer under parallel than under serial processing, because of resource sharing with the other task. Importantly, these predictions are independent of the amount of sharing taking place between the two tasks (assuming SP is not 100:0). That is to say, the AM RT will always be longer under serial than under parallel processing at the 300-ms SOA, regardless of SP, whereas the VV RT will increase under parallel processing with greater SPs (Tombu & Jolicoeur, 2003). This feature of our paradigm is especially advantageous, given that even a minimal amount of capacity sharing—for example, 95 (VV) : 5 (AM)—should be sufficient to produce faster AM RTs under parallel than under serial processing at the 300-ms SOA. Although it has been difficult to observe a strict form of parallel processing—that is, an even distribution of processing resources (Fischer & Hommel, 2012; Lehle & Hübner, 2009; Miller et al., 2009; Plessow, Schade, Kirschbaum, & Fischer, 2012; Tombu & Jolicoeur, 2002, 2005)—previous studies have suggested that such a 95:5 sharing proportion could occur under fixed stimulus and response orders (Oriet, Tombu, & Jolicoeur, 2005; Pashler, 1991), conditions that are common to the present study. Therefore, if capacity can be shared, regardless of whether capacity is evenly distributed between the tasks or biased toward one task, it should be detected with the present paradigm.

Another notable feature of the present task setting is that we chose sensory–motor modality pairings (i.e., audio-manual and visuo-vocal) that are less congruent than the more natural modality pairings (e.g., audio-vocal and visuo-manual). This is because previous studies have shown that natural modality pairings help bypassing the central stage of processing, especially under extensive practice (Hazeltine & Ruthruff, 2006; Levy & Pashler, 2001; Schumacher et al., 2001; Stelzel, Schumacher, Schubert, & D’Esposito, 2005). Given that the goal of the present study was to investigate how information is processed at the central stage, it was incumbent upon us to choose tasks that would engage this stage. To achieve this, we used completely arbitrary stimulus–response mappings and incongruous modality mappings in order to prevent the response selection process at the central stage from benefiting from natural modality mapping. To further ensure that the central stage was not bypassed or eliminated, we included only a moderate amount of practice. If the allocation of central processing resources is truly flexible and under cognitive control, parallel processing should be observed under the present settings.

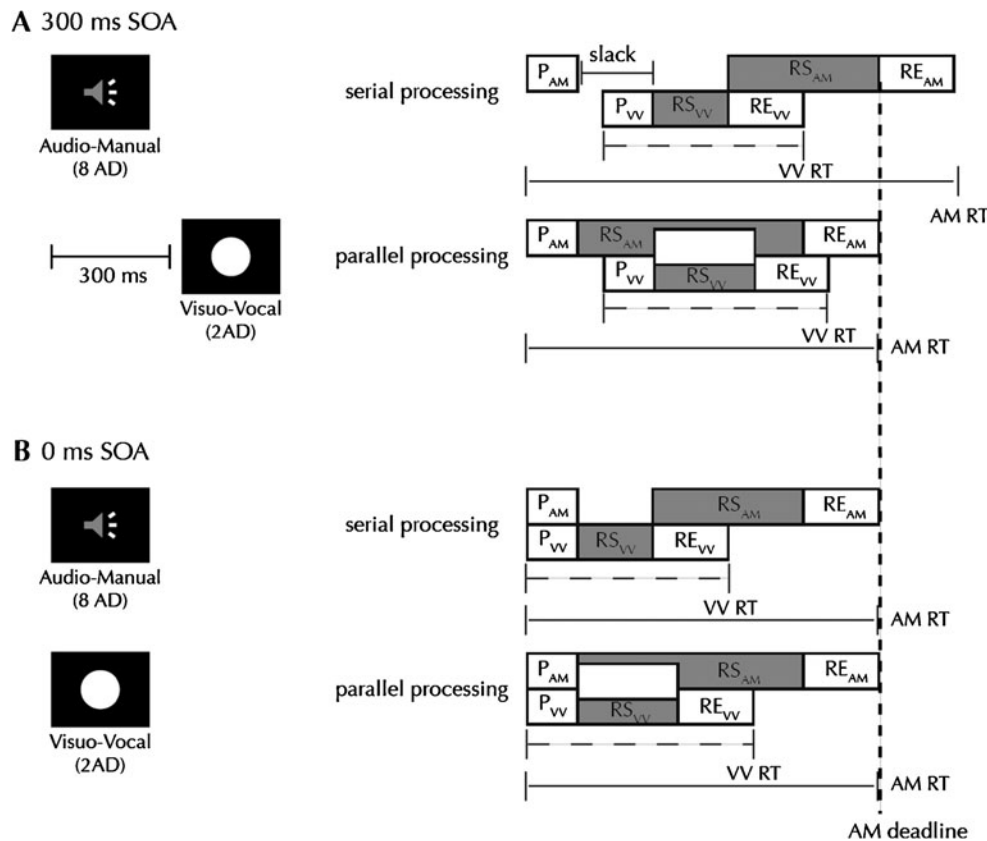


Fig. 2 Predictions of serial and parallel processing modes for Experiment 1a and 1b. **a**At the 300-ms SOA, serial processing induces cognitive slack after perceptual processing of the auditory stimulus, thereby slowing the audio-manual (AM) response relative to parallel processing (and to the AM RT at the 0-ms SOA). In contrast, visuo-vocal (VV) task RTs should be longer with parallel than with serial

processing, because capacity is shared with the AM task. **b**At the 0-ms SOA, parallel and serial processing produce the same AM RT. VV task RTs are expected to be longer in parallel than in serial processing because capacity is shared with the AM task. The AM deadline is set by performance of the AM task at the 0-ms SOA. P, perceptual identification; RS, response selection; RE, response execution

Experiment 1a

Methods

Participants A group of 12 adults (four males, eight females; 18–25 years of age) participated in exchange for financial compensation. The Vanderbilt University Institutional Review Board approved the experimental protocol, and informed consent was obtained from each participant.

Stimuli and apparatus The visual stimulus was a disk (1° of visual angle) presented at the center of the screen and colored either light green or red, with each color requiring a distinct vocal response (either “Bah” or “Koe”). The auditory stimulus was one of eight distinct sounds (Dux et al., 2006), each of which was assigned to a distinct finger keypress.

Design and procedure In each trial two tasks, an 8AD AM task and a 2AD VV task, were presented. There were two SOA conditions (0 and 300 ms), which were blocked. In the 0-ms SOA, the visual and auditory stimuli were presented for

200 ms at the same time, whereas in the 300-ms SOA, the auditory stimulus was always presented first. Participants were instructed to respond to the visual stimulus first in all trials, even when it followed the auditory stimulus.

Participants were also instructed that they would receive bonus pay if they met several criteria. To receive bonus pay, participants not only had to respond to the visual stimulus first, but also complete the AM task within the predetermined RT deadline (see below). Participants were explicitly instructed that processing the two tasks in parallel should outperform processing them serially, and that financial reward is more likely to be earned under parallel processing. Specifically, they were provided an instruction sheet showing Fig. 2, and were instructed that their only chance to receive bonus pay at the longer SOA was to process each task in parallel, and that they should begin processing each of the two tasks as soon as their stimuli appeared rather than wait for the visual stimulus presentation. If participants performed the AM task correctly and within the predetermined deadline, a feedback screen showing “Excellent!” was presented and a bonus pay of 7.8 cents was given. If they were correct for the AM task but missed the

deadline, the feedback “Too slow” was presented (and no bonus pay was allocated). For incorrect responses, the feedback “Wrong” was presented. No feedback was provided for the VV responses. There were four blocks, each consisting of 32 trials, with two of these blocks containing the 0-ms SOA trials and the two others containing the 300-ms SOA trials. Block order was alternated within individuals, and counterbalanced across individuals.

Immediately prior to the main experimental session, an Estimation session was held to determine the AM RT deadline to be used in the main experiment. This deadline was estimated on the basis of the performance in the 0-ms SOA trials because that performance should be the same in the 300-ms SOA if parallel processing can occur. This session contained six blocks of 32 0-ms SOA trials each. The initial three blocks served as familiarization blocks. The last three blocks were carried out as in the main experiment. The mean RT of the last (6th) block was used as the deadline for the main experimental session.

Results and discussion

The results of Experiment 1a are shown in Fig. 3. Reaction times (RT) and accuracy were entered into 2×2 repeated measure ANOVAs with Task (VV and AM) and SOA (0 and 300 ms) as factors. Given that participants were encouraged to respond as fast and accurately as possible, and financial rewards were contingent on both RT and accuracy, it is important to measure processing efficiency independently of potential criterion shifts due, for example, to a speed–accuracy trade-off. To achieve this, we calculated each task’s adjusted RT, also referred to as “inverse efficiency”, a standard measure that combines RT and accuracy by dividing the former by the latter (Graham et al., 2006; Klemen, Verbruggen, Skelton, & Chambers, 2011; Romei, Driver, Schyns, & Thut, 2011; Spence, Kettenmann, Kobal, & McGlone, 2001; Townsend & Ashby, 1983). Even though this inverse efficiency measure is useful in assessing behavioral performance, it should be noted that we do not exclusively rely upon this measure to draw any conclusion from our data. We first present below the analyses on accuracy and RT data, followed by the analyses on adjusted RT data to further support the claims drawn from the accuracy and RT results.

The accuracy results showed main effects of task, $F(1, 11) = 6.91$, $p < .05$, and SOA, $F(1, 11) = 13.91$, $p < .01$, but no interaction, $p > .2$. The RT results also demonstrated main effects of task, $F(1, 11) = 174.90$, $p < .01$, and SOA, $F(1, 11) = 6.38$, $p < .05$, as well as an interaction between the two factors, $F(1, 11) = 105.5$, $p < .01$.

The main effects of SOA on both accuracy and RT (Fig. 3) imply that completing two tasks under instructed response order (VV first) was more challenging at the 300-ms SOA,

which does not support flexible sharing of central resource between the tasks. Most importantly, the AM RT was significantly longer for the 300-ms than for the 0-ms SOA, $t(11) = 5.76$, $p < .01$, whereas the VV RT did not differ across SOAs, $p > .15$. These results suggest that serial, not parallel, processing was adopted because serial processing would slow AM responses due to cognitive slack (Fig. 2).

Interestingly, the observed difference in the AM task RT across SOAs was somewhat smaller (208 ms), $t(11) = 2.53$, $p < .05$, than was predicted by the serial model (300 ms, the duration of the SOA). This smaller-than-expected difference was likely due to participants’ speeding up the AM task in the 300-ms SOA condition to meet the deadline. The fact that AM task accuracy was lower at the 300-ms SOA (paired t -test, $p < .01$) is consistent with this suggestion, as faster responses are more likely to be incorrect. Indeed, when one adjusts RTs for accuracy, the observed difference across SOAs (380 ms) was not significantly different from the predicted value of 300 ms, $p > .27$. By contrast, there was no difference in VV adjusted RT (14 ms) across SOAs, $p > .86$. The invariant VV RT performance across SOAs is notable because it further supports the serial bottleneck model: Differences in VV RTs across SOAs could imply that the two central processes overlapped to different extents across SOAs, which would complicate the interpretation of the present results. However, the fact that only the AM performance was slowed by the predicted amount at the 300-ms SOA suggests that the central stages of the two tasks overlapped completely at both SOAs, and hence that the two response selections took place serially as illustrated in Fig. 2.

The results of the present experiment show that participants failed to flexibly allocate central resource to perform two tasks in parallel; behavioral performance suffered when they were required to respond first to the second-presented stimulus. Furthermore, when RT is adjusted to take into account accuracy, the pattern of behavioral performance fitted well with the prediction derived from the serial bottleneck model.

Experiment 1b

It is conceivable that participants’ failure to perform the AM RT task under the RT deadline at the 300-ms SOA was a result of them simply giving up on meeting these stringent task requirements at that SOA. To examine this possibility, we ran another experiment that included a more lenient AM RT deadline and that randomly intermixed the two SOAs in the experimental session.

Method

The methods were identical to those of Experiment 1a, except for the following modifications. Eight participants

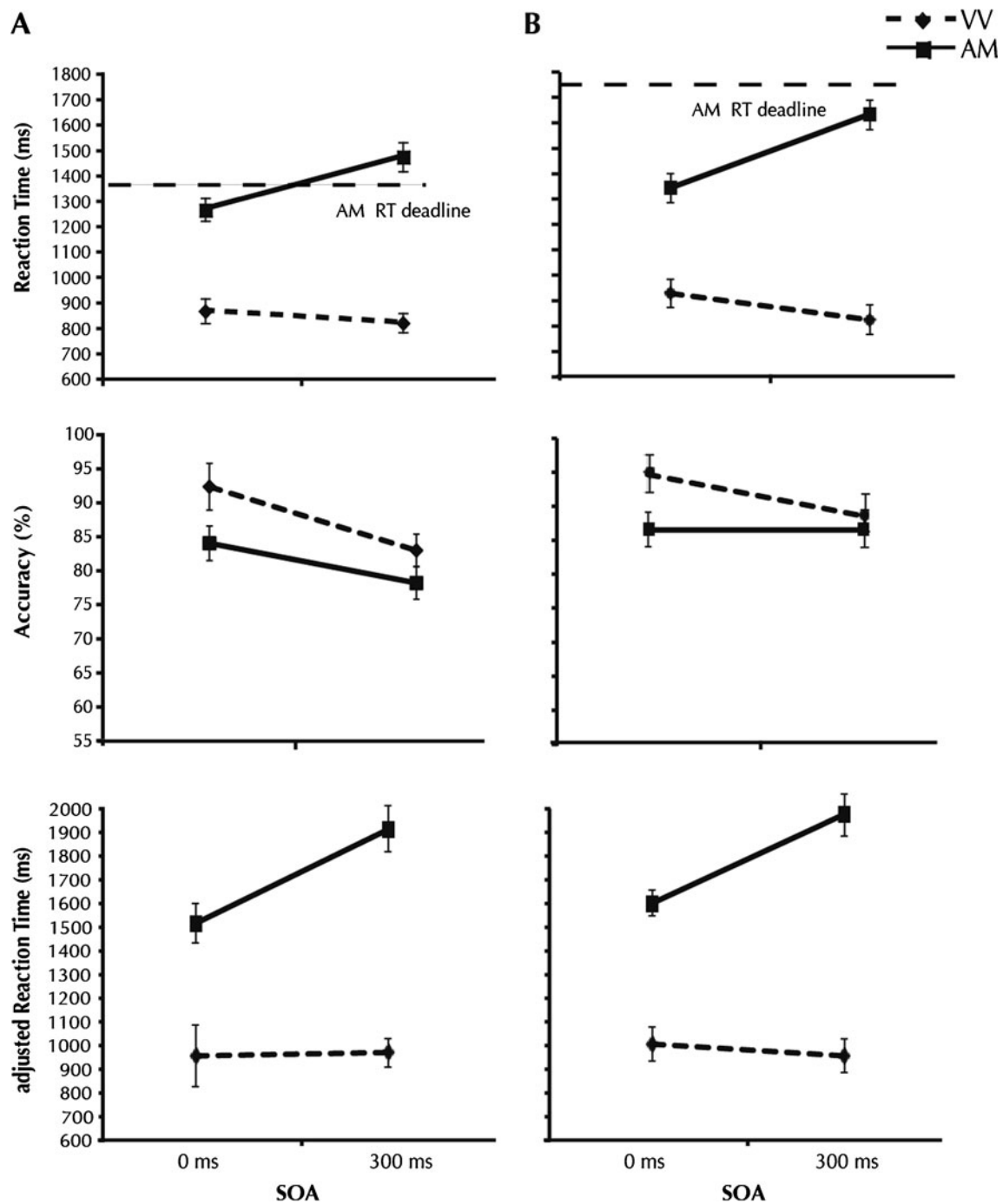


Fig. 3 Results of Experiments 1a(a) and 1b(b). Error bars represent standard errors

(five males, three females; 18–25 years of age) took part, and the deadline estimation session included only three blocks of 0-ms SOA trials without feedback.

Results and discussion

Shortening of the deadline estimation session yielded a much longer deadline (1,750 ms) than had been used in the previous experiment (1,380 ms, $p < .01$). Correspondingly, the

proportion of correct AM responses meeting the deadline increased between experiments from 54 % to 66 % at the 0-ms SOA, and from 30 % to 51 % at the 300-ms SOA, $ps < .05$. Nevertheless, the AM RT for Experiment 1b was still longer at the 300-ms SOA (1,625 vs. 1,335 ms for the 0- vs. 300-ms SOAs), $t(7) = 10.01$, $p < .01$. Furthermore, the difference in AM accuracy between SOAs was eliminated (84.3 % and 84.4 % for the 0- and 300-ms SOAs, $p > .97$), and RT difference increased to 288 ms, which was similar to

the predicted value, $p > .7$.¹ The significant difference in AM adjusted RTs (358 ms, $p < .01$) was also similar to the predicted 300-ms difference, $p > .5$. The latter two findings are consistent with the suggestion that the smaller-than-expected difference in AM task RT across SOAs noted in Experiment 1a was due to subjects rushing their response under the stricter RT deadline, for the expected AM RT difference is obtained when the RT deadline was relaxed. More generally, the results of Experiment 1b suggest that the findings of Experiment 1a were not a consequence of participants giving up on the task due to the strict RT criterion, as similar results were obtained with a more lenient RT deadline.

Taken together, Experiments 1a and 1b suggest that participants could not flexibly allocate resources to each task for optimal performance; when they were required to respond first to the stimulus presented later (300-ms SOA), significant behavioral costs were observed in both RT and accuracy. These behavioral costs at the 300-ms SOA suggest that serial, rather than parallel, processing was adopted. Furthermore, when RT data were adjusted by considering accuracy, or when participants are not trading off accuracy for faster reaction time, the RT pattern fits well with the prediction derived from the serial bottleneck model.

Experiment 2

As we mentioned in the introduction, it has been argued that serial processing only takes place when there has not been sufficient practice with the task conditions to develop the appropriate cognitive control to perform two tasks in parallel (e.g., Schumacher et al., 2001). Even though this assumption is not accepted by all proponents of parallel processing models (Miller et al., 2009; Tombu & Jolicœur, 2003), we considered whether resource sharing did not occur in Experiments 1a and 1b because participants were not sufficiently trained. In the present experiment, we assessed this possibility by increasing the number of trials that participants performed to be equivalent to or larger than the numbers in the studies that have argued for graded sharing (Carrier & Pashler, 1995; Tombu & Jolicœur, 2005). Importantly, whereas this experiment had a

training regimen that was on a par with previous studies of parallel processing, the total number of trials that participants performed was insufficient to completely eliminate dual-task costs, so as to ensure that the response selection stage would not be bypassed through task automatization (Maquestiaux et al., 2008).

Method

The methods were identical to those of Experiment 1a, except for the following modifications. Nine adults (four males, five females; 18–25 years of age) participated in exchange for financial compensation. The number of experimental trials was increased from four to ten blocks of 32 trials (blocked 0-ms and 300-ms SOAs in alternate order, yielding five experimental sessions).

Results and discussions

The results of Experiment 2 are shown in Fig. 4. Accuracy, RTs, and adjusted RTs were entered into $2 \times 2 \times 5$ repeated measure ANOVAs with Task (VV and AM), SOA (0 and 300 ms), and Session (1, 2, 3, 4, and 5) as factors.

The accuracy results showed a main effect of SOA, $F(1, 8) = 15.54$, $p < .01$, consistent with the results of Experiment 1a and 1b. However, contrary to Experiment 1a and 1b, the interaction between task and SOA was significant, $F(1, 8) = 6.31$, $p < .05$. This interaction was due to VV accuracy being significantly lower at the 300-ms SOA (74.5 %) than at the 0-ms (87.8 %) SOA, $t(8) = 3.90$, $p < .01$, whereas we found no difference in AM accuracy (83.2 % for the 0-ms SOA, 79.6 % for the 300-ms SOA), $p > .16$. No other main effect or interaction was significant, $ps > .24$.

The RT results demonstrated a main effect of task, $F(1, 8) = 222.1$, $p < .01$, as well as a main effect of session, $F(4, 32) = 3.72$, $p < .05$. We also observed significant interactions between task and SOA, $F(1, 8) = 57.65$, $p < .01$, and between SOA and session, $F(4, 32) = 6.57$, $p < .01$. No other main effect or interaction was significant, $ps > .29$.

The most important aspect of these results is the difference in AM RTs across SOAs. The AM response was significantly slower for the 300-ms SOA than for the 0-ms SOA, $t(8) = 5.03$, $p < .01$, consistent with the findings of Experiment 1a and 1b, but inconsistent with the graded-sharing model. However, that difference (137 ms) was smaller than the 300 ms ($p < .01$) predicted by the serial bottleneck model. Interestingly, a highly significant difference in VV RTs also emerged (118 ms), $t(8) = 3.61$, $p < .01$, and the faster VV response at the 300-ms SOA was accompanied by a significantly lower accuracy at this SOA, as evidenced by the significant interaction between task and SOA on accuracy (see above). We hypothesized that the smaller-than-300-ms AM RT difference was due to participants speeding up the VV

¹ The results of this experiment also ruled out a potential “practice” account for the AM RT difference between SOAs: Because only the 0-ms SOA condition was used in the estimation sessions of this experiment and Experiment 2, faster RTs at the 0-ms than at the 300-ms SOA in the main session may have been caused by greater exposure to the former SOA trials. However, were this the case, a practice benefit should also have been observed for the VV task, but that was not the case for either experiment. Furthermore, whereas the estimation session of the present experiment was half as long as that of queryExperiment 1a, it produced a bigger RT difference between the 0-ms and 300-ms SOAs in the experimental session, a finding that is inconsistent with a practice effect.

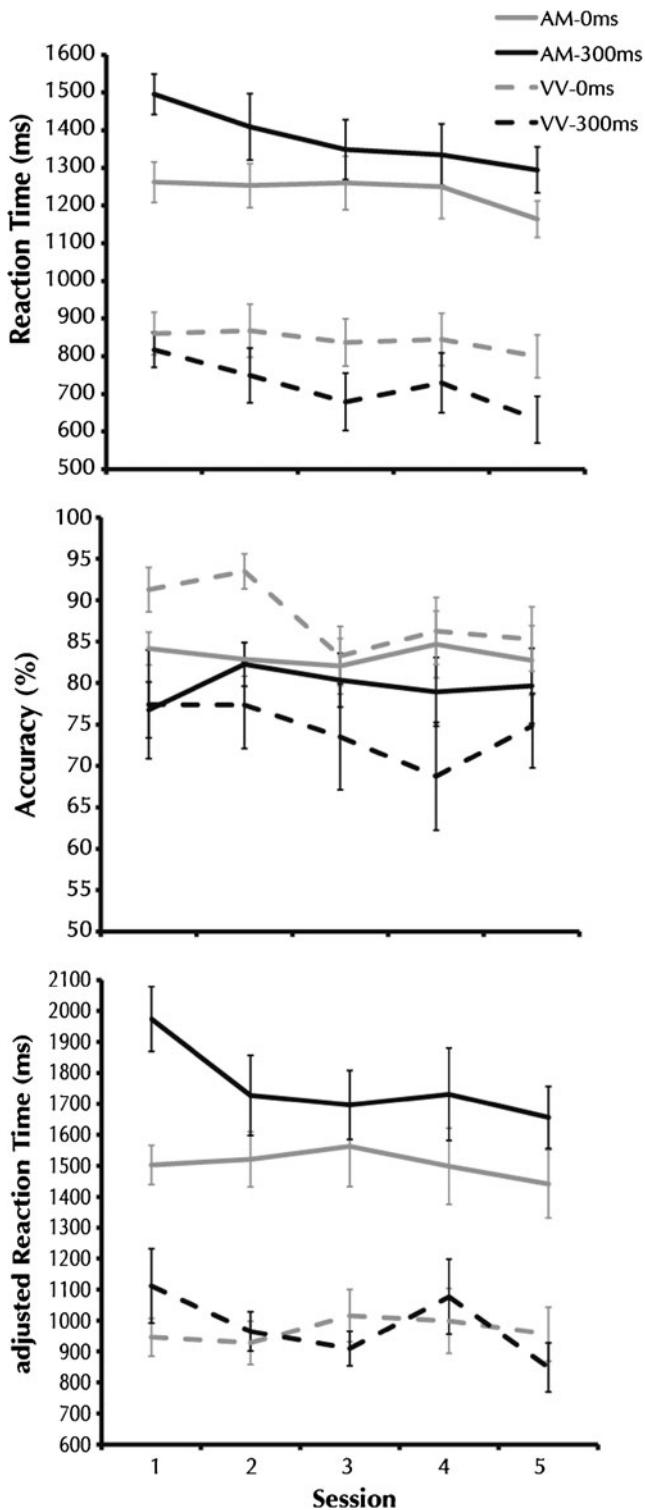


Fig. 4 Results of Experiment 2. Error bars represent standard errors

response in order to meet the RT deadline for the AM task. When they noticed that meeting the response deadline was harder at the 300-ms SOA, participants may have sped up the VV response, curtailing its processing and sacrificing its accuracy. As a result of faster VV processing, participants could

process the AM task sooner, thereby reducing its RT. Consistent with this proposition, the sum of the AM RT and VV RT differences between SOAs was not different from the predicted value of 300 ms ($254 \text{ ms}, p > .21$). As a whole, these results are consistent with the serial bottleneck model of information processing.

The hypothesis that serial, not parallel, processing occurred was further supported when the present RT data were separately examined for each session and task. As is shown in Fig. 4, the AM RT difference across SOAs remained significant across all sessions ($ps < .01$). Importantly, although practice decreased the AM RT difference across SOAs, it only did so from the 1st to the 3rd session, $p < .01$, with no further decrease thereafter (130-ms AM RT difference in the final session). Consistent with the explanation that the smaller-than-300-ms AM RT difference between SOAs resulted from the speeding up of VV RTs, a VV RT difference of 43 ms ($p = .30$) in the 1st session increased to 157 ms ($p < .01$) by the 3rd session, when it leveled off, mirroring the AM RT pattern. Moreover, not only did we find a positive correlation between the VV and AM RTs in all sessions (r^2 ranged from .21 to .80, all $ps < .05$), but, critically, the sum of the AM RT and VV RT differences between SOAs was not different from the predicted value of 300 ms even by the last training session (297 ms, $t < 1$).

Finally, adjusted RTs were analyzed in the same way as accuracy and RTs. We observed main effects of task, $F(1, 8) = 282.1, p < .01$, and SOA, $F(1, 8) = 15.89, p < .01$, as well as interactions between task and SOA, $F(1, 8) = 13.39, p < .01$, and between SOA and session, $F(1, 8) = 3.48, p < .05$. The adjusted VV RT did not differ across SOAs, $t < 1$. By contrast, the AM response at the 300-ms SOA was significantly slower, by 235 ms, $t(8) = 3.75, p < .01$, which was indistinguishable from the predicted value of 300 ms, $p > .30$. Taken together, these results strongly suggest that even under practice conditions, serial processing still persisted.

General discussion

In the present study, we tested whether people could flexibly allocate central capacity resources across tasks in order to optimize their behavioral performance. We did so by creating a situation in which task requirements could only be fulfilled successfully by adopting parallel processing, by financially rewarding the adoption of parallel processing, and by making this mode of processing more efficient (i.e., having a lower sum of task RTs) than serial processing. However, despite these favorable conditions, participants failed to adopt parallel processing, and instead performed response selections serially. This lack of evidence for parallel processing could not be attributed to the possibility that participants

gave up following the task instruction due to the strict deadline (Exp. 1a), nor could it be attributed to an insufficient amount of practice to develop parallel processing (Exp. 2). Instead, the results of the present experiments suggest that the serial postponement of the second response in dual-task is obligatory, reflecting the intrinsic nature of the central capacity limit.

One might argue that the present dual-task structure—which involved responding first to the stimulus that was presented last, pairing two tasks that have substantially different task demands (2AD vs. 8AD), and using incongruous sensory–motor pairings (visuo-vocal and audio-manual tasks, instead of audio-vocal and visuo-manual ones)—prohibited the adoption of parallel processing. For example, some response modality pairings can be more conducive to parallel processing than others (Hazeltine & Ruthruff, 2006; Levy & Pashler, 2001), and increasing task demands may incite serial processing (Fischer, Miller, & Schubert, 2007; Luria & Meiran, 2005; but see also Tombu & Jolicoeur, 2002). Given that such factors in task settings can influence the adoption of specific task strategies (Fischer et al., 2007), it should not be interpreted from the present findings that parallel processing is not possible. However, if shared allocation of processing resources is so constrained that parallel processing cannot be implemented even when it would yield better task performance than serial processing, we are compelled to conclude that the allocation of central processing resources is not as flexible as has been asserted in certain cognitive models of parallel processing.

A related issue concerning the choice of response modality pairing is that it might induce “crosstalk” between tasks that could affect dual-task performance above and beyond any costs associated with limitations at a central stage of information processing (Logan & Schulkind, 2000; Stelzel & Schubert, 2011). Such crosstalk might be particularly costly for task pairings that are not natural (Stelzel & Schubert, 2011). However, such crosstalk cannot easily explain why the AM task performance at the 300-ms SOA was worse than that at the 0-ms SOA, because such crosstalk should be equivalent at both SOAs. A more critical issue could be that participants adopted a serial processing mode in order to reduce such crosstalk. However, in the present paradigm, sharing capacity and processing the two tasks in parallel would be the only way to meet the criteria to get rewarded, for which our participants were highly motivated. Hence, any crosstalk that might have occurred in our tasks should not affect the principal findings and interpretations of the present study.

It is worth pointing out that, in addition to the incongruous sensory–motor mapping, the strict constraint on the order of stimulus presentation and response has also been known to favor serial processing (Meyer & Kieras, 1999). Indeed, many dual-task studies have presented two tasks in a random

order and used unconstrained response orders, to discourage serial processing and facilitate “perfect time sharing” between two tasks. Although a number of studies have implied that such randomization of stimulus and response orders is important for observing reductions in dual-task costs (Liepelt, Fischer, Frensch, & Schubert, 2011; Schumacher et al., 2001), in the present study we did not pursue the issue of which factors are important to reduce or eliminate dual-task costs. As we mentioned in the introduction, the present study was aimed at unraveling how multiple response selections proceed at the capacity-limited central stage. To address this issue with our novel paradigm, the constraint on the stimulus presentation and response order was essential. Moreover, it has been argued that capacity sharing can occur even under such experimental constraints (Oriet et al., 2005; Pashler, 1991).

The graded-sharing model was developed to accommodate some empirical findings that did not fit well with earlier parallel models (McLeod, 1977). With the assumptions that capacity is limited only at the central stage and that it is possible to reallocate capacity instantaneously, without cost, when a given task is completed (Kahneman, 1973), the model has the power to explain nearly all patterns of extant PRP data and, as is common for resources theories (Navon, 1984), has been hard to falsify. Yet its predictions have not been easy to empirically support (Tombu & Jolicoeur, 2002), either, and some of these findings are open to alternative interpretations (Oriet et al., 2005; Tombu & Jolicoeur, 2005). These difficulties have been primarily attributed to the fact that parallel processing is generally less favored because it might be more effortful or less efficient than serial processing (Miller et al., 2009; Tombu & Jolicoeur, 2002, 2005), although these claims have been challenged (Lehle, Steinhauser, & Hübner, 2009). Indeed, one study has suggested that parallel processing could be favored because it reduces mental effort, by sacrificing processing efficiency (Plessow et al., 2012). However, here we showed that participants fail to perform tasks in parallel even when that is the only means by which they can successfully perform the task.

It should be acknowledged that we did not directly test other models of parallel processing than the graded-sharing model. Hence, the present results do not rule out the possibility that some other form of parallel processing might take place. In particular, it is possible that an extensive amount of practice might allow one to develop parallel processing (Meyer & Kieras, 1997; Schumacher et al., 2001), although that viewpoint has been challenged (Miller et al., 2009; Ruthruff et al., 2009). In addition to practice, specific combinations of response modality mappings can also aid in the emergence of parallel processing (Levy & Pashler, 2001). Thus, it is possible that parallel processing could have emerged in the present study had we used more natural modality pairings and exposed participants to extensive

practice with the tasks. Under such circumstances, the central processing stage is more likely to be bypassed or eliminated, allowing simultaneous performance of two tasks (Dux et al., 2009; Hazeltine & Ruthruff, 2006; Levy & Pashler, 2001; Schumacher et al., 2001). However, the graded-sharing model is the one that accommodates most of the extant PRP data, even under conditions that do not involve much practice, and irrespective of modality pairings. More importantly, this model shares the core assumption that many other cognitive-control models of the PRP posit: the flexible control of resource allocation. Thus, at the very least, the present finding that the allocation of central resources recruited for response selection is not flexible constrains cognitive-control models of human dual-task performance. Indeed, it points to this central stage in human information processing as being primarily limited by a serial bottleneck.

Author note This work was supported by NIMH Grant No. RO1 MH70776 to R.M., and by Grant No. P30-EY008126 to the Vanderbilt Vision Research Center. We thank Mike Tombu for valuable comments on the manuscript.

References

- Carrier, L. M., & Pashler, H. (1995). Attentional limits in memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 1339–1348.
- Dux, P. E., Ivanoff, J., Asplund, C. L., & Marois, R. (2006). Isolation of a central bottleneck of information processing with time-resolved fMRI. *Neuron*, *52*, 1109–1120.
- Dux, P. E., Tombu, M. N., Harrison, S., Rogers, B. P., Tong, F., & Marois, R. (2009). Training improves multitasking performance by increasing the speed of information processing in human prefrontal cortex. *Neuron*, *63*, 127–138. doi:10.1016/j.neuron.2009.06.005
- Fischer, R., & Hommel, B. (2012). Deep thinking increases task-set shielding and reduces shifting flexibility in dual-task performance. *Cognition*, *123*, 303–307. doi:10.1016/j.cognition.2011.11.015
- Fischer, R., Miller, J., & Schubert, T. (2007). Evidence for parallel semantic memory retrieval in dual tasks. *Memory & Cognition*, *35*, 1685–1699. doi:10.3758/BF03193502
- Graham, K. S., Scabill, V. L., Hornberger, M., Barense, M. D., Lee, A. C., Bussey, T. J., & Saksida, L. M. (2006). Abnormal categorization and perceptual learning in patients with hippocampal damage. *Journal of Neuroscience*, *26*, 7547–7554. doi:10.1523/JNEUROSCI.1535-06.2006
- Hazeltine, E., & Ruthruff, E. (2006). Modality pairing effects and the response selection bottleneck. *Psychological Research*, *70*, 504–513.
- Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs: Prentice Hall.
- Kamienkowski, J. E., Pashler, H., Dehaene, S., & Sigman, M. (2011). Effects of practice on task architecture: Combined evidence from interference experiments and random-walk models of decision making. *Cognition*, *119*, 81–95.
- Klemen, J., Verbruggen, F., Skelton, C., & Chambers, C. D. (2011). Enhancement of perceptual representations by endogenous attention biases competition in response selection. *Attention, Perception, & Psychophysics*, *73*, 2514–2527. doi:10.3758/s13414-011-0188-5
- Lehle, C., & Hübner, R. (2009). Strategic capacity sharing between two tasks: Evidence from tasks with the same and with different task sets. *Psychological Research*, *73*, 707–726. doi:10.1007/s00426-008-0162-6
- Lehle, C., Steinhäuser, M., & Hübner, R. (2009). Serial or parallel processing in dual tasks: What is more effortful? *Psychophysiology*, *46*, 502–509.
- Levy, J., & Pashler, H. (2001). Is dual-task slowing instruction dependent? *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 862–869.
- Liepert, R., Fischer, R., Frensch, P. A., & Schubert, T. (2011). Practice-related reduction of dual-task costs under conditions of a manual-pedal response combination. *Journal of Cognitive Psychology*, *23*, 29–44.
- Logan, G. D., & Schulkind, M. D. (2000). Parallel memory retrieval in dual-task situations: I. Semantic memory. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 1072–1090. doi:10.1037/0096-1523.26.3.1072
- Luria, R., & Meiran, N. (2005). Increased control demand results in serial processing: Evidence from dual-task performance. *Psychological Science*, *16*, 833–840. doi:10.1111/j.1467-9280.2005.01622.x3
- Maquestiaux, F., Laguë-Beauvais, M., Bherer, L., & Ruthruff, E. (2008). Bypassing the central bottleneck after single-task practice in the psychological refractory period paradigm: Evidence for task automatization and greedy resource recruitment. *Memory & Cognition*, *36*, 1262–1282. doi:10.3758/MC.36.7.1262
- McLeod, P. (1977). Parallel processing and the psychological refractory period. *Acta Psychologica*, *41*, 381–391.
- Meyer, D., & Kieras, D. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part 2. Accounts of a psychological refractory-period phenomenon. *Psychological Review*, *104*, 749–791. doi:10.1037/0033-295X.104.4.749
- Meyer, D. E., & Kieras, D. E. (1999). Precis to a practical unified theory of cognition and action: Some lessons from EPIC computational models of human multiple-task performance. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII: Cognitive regulation of performance. Interaction of theory and application* (pp. 15–88). Cambridge: MIT Press.
- Miller, J., Ulrich, R., & Rolke, B. (2009). On the optimality of serial and parallel processing in the psychological refractory period paradigm: Effects of the distribution of stimulus onset asynchronies. *Cognitive Psychology*, *58*, 273–310. doi:10.1016/j.cogpsych.2006.08.003
- Navon, D. (1984). Resources—A theoretical soup stone? *Psychological Review*, *91*, 216–234.
- Navon, D., & Miller, J. (2002). Queuing or sharing? A critical evaluation of the single-bottleneck notion. *Cognitive Psychology*, *44*, 193–251.
- Oriet, C., Tombu, M., & Jolicœur, P. (2005). Symbolic distance affects two processing loci in the number comparison task. *Memory & Cognition*, *33*, 913–926. doi:10.3758/BF03193085
- Pashler, H. (1984). Processing stages in overlapping tasks: Evidence for a central bottleneck. *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 358–377.
- Pashler, H. (1991). Shifting visual attention and selecting motor responses: Distinct attentional mechanisms. *Journal of Experimental Psychology: Human Perception and Performance*, *17*, 1023–1040.
- Pashler, H. (1994a). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin*, *116*, 220–244. doi:10.1037/0033-2909.116.2.220

- Pashler, H. (1994b). Graded capacity-sharing in dual-task interference? *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 330–342. doi:10.1037/0096-1523.20.2.330
- Plessow, F., Schade, S., Kirschbaum, C., & Fischer, R. (2012). Better not to deal with two tasks at the same time when stressed? Acute psychosocial stress reduces task shielding in dual-task performance. *Cognitive, Affective & Behavioral Neuroscience*, *12*, 557–570. doi:10.3758/s13415-012-0098-6
- Romei, V., Driver, J., Schyngs, P. G., & Thut, G. (2011). Rhythmic TMS over parietal cortex links distinct brain frequencies to global versus local visual processing. *Current Biology*, *21*, 334–337.
- Ruthruff, E., Hazeltine, E., & Remington, R. W. (2006). What causes residual dual-task interference after practice? *Psychological Research*, *70*, 494–503.
- Ruthruff, E., Johnston, J. C., & Remington, R. W. (2009). How strategic is the central bottleneck: Can it be overcome by trying harder? *Journal of Experimental Psychology: Human Perception and Performance*, *35*, 1368–1384.
- Ruthruff, E., Johnston, J. C., & Van Selst, M. (2001). Why practice reduces dual-task interference. *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 3–21. doi:10.1037/0096-1523.27.1.3
- Schumacher, E. H., Seymour, T. L., Glass, J. M., Fencsik, D. E., Lauber, E. J., Kieras, D. E., & Meyer, D. E. (2001). Virtually perfect time sharing in dual-task performance: Uncorking the central cognitive bottleneck. *Psychological Science*, *12*, 101–108. doi:10.1111/1467-9280.00318
- Spence, C., Kettenmann, B., Kopal, G., & McGlone, F. P. (2001). Shared attentional resources for processing visual and chemosensory information. *Quarterly Journal of Experimental Psychology*, *54*, 775–783.
- Stelzel, C., & Schubert, T. (2011). Interference effects of stimulus–response modality pairings in dual tasks and their robustness. *Psychological Research*, *75*, 476–490.
- Stelzel, C., Schumacher, E. H., Schubert, T., & D’Esposito, M. (2005). The neural effect of stimulus–response modality compatibility on dual-task performance: An fMRI study. *Psychological Research*, *70*, 514–525. doi:10.1007/s00426-005-0013-7
- Tombu, M., & Jolicoeur, P. (2002). All-or-none bottleneck versus capacity sharing accounts of the psychological refractory period phenomenon. *Psychological Research*, *66*, 274–286.
- Tombu, M., & Jolicoeur, P. (2003). A central capacity sharing model of dual-task performance. *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 3–18. doi:10.1037/0096-1523.29.1.3
- Tombu, M., & Jolicoeur, P. (2004). Virtually no evidence for virtually perfect time-sharing. *Journal of Experimental Psychology: Human Perception and Performance*, *30*, 795–810.
- Tombu, M., & Jolicoeur, P. (2005). Testing the predictions of the central capacity sharing model. *Journal of Experimental Psychology: Human Perception and Performance*, *31*, 790–802.
- Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modelling of 1240 elementary psychological processes*. Cambridge: Cambridge University Press.
- 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.