Journal of Experimental Psychology:
Learning, Memory, and Cognition
2003, Vol. 29, No. 3, 373–388

0278-7393/03/$12.00 DOI: 10.1037/0278-7393.29.3.373

Practice Effects on Two Memory Retrievals From a Single Cue

Randall S. Nino and Timothy C. Rickard
University of California, San Diego

The effects of practice on 2 retrievals from a single cue were investigated. In Experiment 1, participants were given extended single-task practice and were then tested on a dual memory retrieval task. Performance was consistent with a sequential retrieval model proposed by T. C. Rickard and H. Pashler (2003). In Experiment 2, participants practiced both single- and dual-retrieval tasks extensively. Initially, data from all participants indicated sequential retrieval. However, participants who grouped the 2 response outputs were eventually able to perform the dual task with a latency that approached the prediction of a parallel race model. Models that assume a transition from sequential to parallel retrieval with practice, along with other models that assume an immutable retrieval bottleneck at all practice levels, are considered.

The question of whether memory retrieval can be executed in parallel with other cognitive processes, including other retrievals, is currently an active topic of investigation. In several recent articles, theorists have advanced the idea that two or more memory instances can be retrieved in parallel (Logan, 1988, 1992; Nosofsky & Palmeri, 1997; Palmeri, 1997, 1999). Other researchers have explicitly suggested that two retrievals can proceed in parallel from a common cue (Rohrer, Pashler, & Etchegaray, 1998; Ross & Anderson, 1981; Wenger, 1999). However, a recent study conducted by Rickard and Pashler (2003) suggests otherwise. In their first experiment, participants learned two responses for each of 10 visually presented color names. In the vocal task, participants learned to speak a unique single-digit number when presented with the cue. In the keypress task, participants learned to press one of two keys as a response. In their second experiment, there were eight cue words, eight unique vocal-digit responses, and eight unique keypress responses. In each experiment, after learning the two tasks individually, participants entered a test phase consisting of interleaved blocks of the vocal task, the keypress task, and a new dual task in which they both spoke the number and pressed the key when the cue appeared. Because the two responses involved two different modalities, there should have been minimal interference at the response output level of processing (Pashler & Christian, 1996; Fagot & Pashler, 1992; Schubert, 1999). Thus, any bottleneck in performance that may exist must reflect the retrieval stage of performance (Schubert, 1999).

The mean time to complete the dual task in Rickard and Pashler's (2003) study was close to the sum of the vocal and keypress single-task retrieval times, with the exception of the first test block, on which the dual-task times were somewhat slower than that sum. In contrast, the mean time to complete the dual task was several hundred milliseconds above the prediction of an unlimited capacity parallel retrieval model (often termed a race model). These findings, along with supporting analyses based on dual-task response ordering and dual-task response time (RT) correlations, indicated that participants retrieved first one response, and then the other, sequentially. These results are consistent with other recent research in related domains (e.g., Carrier & Pashler, 1995; Craik, Govoni, Naveh-Benjamin, & Anderson, 1996), suggesting a central processing bottleneck that allows only one memory retrieval to be completed at a time, at least at relatively low levels of practice.

Rickard and Pashler (2003) observed two distinct patterns across participants in the dual-task interresponse interval, or IRI (i.e., the difference between RT2, the time to complete both tasks, as measured from the onset of the stimulus, and RT1, the time to complete the first task). A number of participants had very short IRIs, averaging approximately 200 ms or less. Rickard and Pashler used the term *response groupers* to refer to these participants. They inferred that these individuals adopted a strategy of waiting until both retrievals were completed before executing the two responses. On the other hand, about half of the participants had very long IRIs between responses, always greater than 344 ms and averaging 554 ms. These participants, whom Rickard and Pashler termed *nongroupers,* seemed to retrieve one response first and then execute that response immediately before retrieving and executing the second response. For both groupers and nongroupers, the RT2 was approximately the sum of the two single-task RTs, indicating that retrieval is sequential regardless of response strategy.

This article explores whether the Rickard and Pashler (2003) results continue to hold even after participants are given extensive practice on the single tasks and dual tasks. A core theoretical issue here is whether automatization of the retrieval skill will eliminate the bottleneck, allowing the two retrievals to be executed in parallel. It has been demonstrated in related task domains that participants often use controlled mnemonic strategies early in retrieval practice but then make a transition to a more automatic, direct retrieval process later (e.g., Crutcher & Ericsson, 2000; Richardson, 1998). It could well be that once this transition has been made, retrieval of the two responses can proceed in parallel. To explore this possibility in the first experiment, the two single

---

Randall S. Nino and Timothy C. Rickard, Department of Psychology, University of California, San Diego.

This research was supported by National Institute of Mental Health Grant R29 MH58202-01A1 to Timothy C. Rickard.

Correspondence concerning this article should be addressed to Timothy C. Rickard, University of California, San Diego, Department of Psychology, 0109, 9500 Gilman Drive, La Jolla, California 92093-0109. E-mail: trickard@ucsd.edu

tasks were practiced extensively, but independently, before the first dual-task block. The dual task phase will provide insight into whether two highly practiced and independent retrievals from a single cue can be executed in parallel. The goal of the second experiment was to determine whether the model that best describes performance when the two tasks are independent continues to describe performance after dual-task practice, when the tasks are no longer guaranteed to be independent. We also sought to determine whether participants' response strategy (i.e., grouping vs. nongrouping) has any influence on the effects of extended dual-task practice.

Quantitative Models

Predictions of both a sequential model and a race model can be derived directly from the single-task data, requiring no free parameters, as described by Rickard and Pashler (2003). Both of these models assume that perception, memory retrieval, and motor components of processing (*p*, *r*, and *m*, respectively) are stochastically independent serial processes.[1]

The Race Model

According to the race model, there is first a single perceptual event, *p*, for the stimulus, along with a dual-task preparation stage, prep, followed by parallel retrieval and motor execution for the two tasks. Assuming that each response is executed as soon as it is available (i.e., assuming that responses are not grouped), this model predicts that, for each dual-task trial:

$$\mu_{\mathrm{RT1}} = \mu_p + \mu_{\mathrm{prep}} + \mu_{[\min(\mathrm{RT}kr+\mathrm{RT}km,\mathrm{RT}vr+\mathrm{RT}vm)]} \tag{1}$$

and

$$\mu_{\mathrm{RT2}} = \mu_p + \mu_{\mathrm{prep}} + \mu_{[\max(\mathrm{RT}kr+\mathrm{RT}km,\mathrm{RT}vr+\mathrm{RT}vm)]}, \tag{2}$$

where μ refers to the population mean of an RT distribution, RT1 is the latency to the first response made on a dual-task trial (regardless of which task is completed first), RT2 is the latency to the second response made on a dual-task trial, RT*kr* + RT*km* is the combined retrieval and motor component latency of a keypress trial, and RT*vr* + RT*vm* is the combined retrieval and motor component latency of a vocal trial. The terms min(RT*kr* + RT*km*, RT*vr* + RT*vm*) and max(RT*kr* + RT*km*, RT*vr* + RT*vm*) refer to the prediction that, if one keypress and one vocal response are drawn randomly from their respective distributions on a dual-task trial, RT1 will correspond to the minimum, or smaller, of the pair, and RT2 will correspond to the maximum, or larger, of the pair. Note that even if this model correctly describes the underlying retrieval processes, μ_{RT1} will exceed the value predicted in Equation 1 if participants group responses. Thus, Equation 1 is applicable as a prediction only when responses are not grouped (for related discussions of grouping effects, see Pashler, 1994; Pashler & Johnston, 1989; Rickard & Pashler, 2003). However, assuming that the grouping strategy induces negligible processing delay, Equation 2 should remain a valid prediction for RT2.

The predictions of Equations 1 and 2 were estimated from the single-task data following Rickard and Pashler (2003). The term μ_{prep} was treated as a constant with a value of 100 ms. A dual-task preparation cost of about 100 ms has been observed in other dual-task domains (see Li & Wright, 2000; Pashler, 1998, p. 274), and the results of Rickard and Pashler suggested that its magnitude might be about 100 ms in our dual retrieval task as well.

Setting aside μ_{prep} for the moment, estimates for the combined latency of the remaining components of Equations 1 and 2 were generated from the single-task data in the following way. Each individual dual-task trial was matched to the trials on the immediately preceding single-task vocal and keypress blocks that had the same cue. For each of these data trios, the smaller and larger of the two single-task RTs were determined. The smaller, or minimum, RT provided an estimate of a single observation drawn from the theoretical RT1 distribution for the dual-task item of that trio, and the larger, or maximum, RT provided a single observation drawn from the theoretical RT2 distribution for that same dual-task item. Consider, for example, the case in which a participant is presented with the cue "red" on a particular dual-task block. A random sample from the theoretical distribution for RT1 under the parallel model can be approximated by simply selecting the faster of the keypress and vocal single RTs for the cue "red" from the immediately preceding single-task test blocks. Analogously, an estimate for RT2 can be obtained by selecting the slower RT. By averaging these minimum and maximum RTs over items and participants, estimates of Equations 1 and 2 were obtained.[2]

Although roughly accurate, these estimates implicitly treat the single perceptual event of the dual task as if it represented two independent perceptual events, one for each task, operating in parallel. Thus, stated as equations, the resulting predictions for μ_{RT1} and μ_{RT2} for each dual-task trial then correspond to:

$$\mu_{\mathrm{RT1}} = \mu_{\mathrm{prep}} + \mu_{[\min(\mathrm{RT}p+\mathrm{RT}km+\mathrm{RT}kr,\mathrm{RT}p+\mathrm{RT}vm+\mathrm{RT}vr)]} \tag{3}$$

and

$$\mu_{\mathrm{RT2}} = \mu_{\mathrm{prep}} + \mu_{[\max(\mathrm{RT}p+\mathrm{RT}km+\mathrm{RT}kr,\mathrm{RT}p+\mathrm{RT}vm+\mathrm{RT}vr)]}. \tag{4}$$

The inclusion of RT*p* in these parallel processing components will tend to bias the estimates of μ_{RT1} and μ_{RT2} to be somewhat smaller than should be the case according to Equations 1 and 2. In the dual-task, RT*p* is a lone task component contributing to both RT1 and RT2. The estimates provided by Equations 3 and 4, however, treat RT*p* as if it were an independent event for each task. To access the magnitude of this bias, we performed the simulations discussed in the Appendix. Under empirically motivated assumptions about the population mean and variance of RT components in the preceding equations, the simulations show that the induced bias is less than about 12 ms for both RT1 and RT2. These effects are small relative to the scale of the observed RTs (about 800 to 2,500 ms) and can safely be ignored in the following analyses without compromising theoretical inference.

---

[1] There is independent support for the assumption of stochastically independent and serially additive perception, retrieval, and motor response stages (see, for example, Scarborough & Sternberg, 1998).

[2] The method of sampling RTs from the single-task block immediately preceding each dual-task block could result in a slight bias, causing the models to overestimate RT1 and RT2. This could occur because there might be an RT speedup effect due to performance of each single-task test block. Given the single-task results of these experiments, however, this bias did not appear to unduly influence the model fits. Use of the single-task blocks after each dual-task block to make the matched predictions yielded predictions similar to the preceding single-task blocks.

The Sequential Retrieval Model

According to the sequential model, the memory retrieval components of processing, μ_{kr} and μ_{vr}, must occur sequentially on dual-task trials, but it allows for either of them to take place in parallel with the motor response execution for the other task. For example, if the keypress retrieval is completed first, the keypress motor output can take place while the memory retrieval for the vocal task is being executed. The model makes no predictions about which task will be completed first on a given trial. Assuming that each response is executed as soon as it is available, the sequential prediction for μ_{RT1} for each dual-task trial is

$$\mu_{\text{RT1}} = \mu_p + \mu_{\text{prep}} + (j)(\mu_{kr} + \mu_{km}) + (1 - j)(\mu_{vr} + \mu_{vm}). \quad (5)$$

The value j is either 1, if the keypress task is completed first, or 0, if the vocal task is completed first. The value of j was determined empirically for each prediction of each dual-task trial by simply noting which of the two responses was given first on that trial. Consider, for example, a dual-task trial on which the keypress task is completed first. In this case, the RT1 prediction reduces to $\mu_{\text{RT1}} = \mu_p + \mu_{\text{prep}} + \mu_{kr} + \mu_{km}$, which is simply μ_{prep} combined with the observed single-task keypress RT on the preceding single-task block. Note that even if the sequential model correctly describes the underlying retrieval process, μ_{RT1} will exceed the prediction of Equation 5 for any participant who adopts a response grouping strategy.

The sequential prediction for RT2 on each dual-task trial is

$$\mu_{\text{RT2}} = \mu_p + \mu_{\text{prep}} + \mu_{kr} + \mu_{vr} - [(1 - j)\mu_{km} + (j)\mu_{vm}]. \quad (6)$$

If the keypress response is retrieved first, then the motor component of the keypress RT, μ_{km}, can, according to this model, take place in parallel with the retrieval of the vocal task, and hence that component can be subtracted from μ_{RT2}. In this case, the term $[(j)\mu_{km} + (1 - j)\mu_{vm}]$ reduces to μ_{vm}. The analogous logic holds for trials on which the vocal task is retrieved first.

As for the race model, μ_{prep} was assumed to be 100 ms. Estimates for the combined latency of the remaining components of Equation 6 were generated from the single-task data in a manner analogous to the case of the race model. First, each individual dual-task trial was matched to the trials on the immediately preceding single-task vocal and keypress blocks that had the same cue. The single task vocal and keypress task RTs were then summed. By averaging these sums over items and participants, just as for the RT2 of the dual task, an estimate of $\mu_p + \mu_{kr} + \mu_{vr} - [(j)\mu_{km} + (1 - j)\mu_{vm}]$ was obtained (i.e., the prediction of Equation 6 minus μ_{prep}). Combining this estimate with μ_{prep} yields an estimate of RT2 as predicted by Equation 6.

Although roughly accurate, the method for estimating RT2 just outlined is slightly biased. To see this, first consider the RT components contained within the predictions based on the single-task data, given by

$$\mu_{\text{RT2}} = \mu_{\text{prep}} + (\mu_p + \mu_{kr} + \mu_{km}) + (\mu_p + \mu_{vr} + \mu_{vm}). \quad (7)$$

In comparison with Equation 6, this equation contains one extra perceptual term, μ_p, and one extra motor response term, either μ_{km} or μ_{vm}, depending on which task was completed first. This resulting bias can be corrected by subtracting an estimate of the combined perceptual and motor component of the task that was completed first on a given dual-task trial.

Pashler and Johnston (1989; see also Pashler, 1994) conducted a serial RT study that provides a reliable estimate of about 200 ms for the sum of perceptual and motor component RTs for tasks involving visual input and keypress output. Note, however, that their data, which involved keypress responses, can be applied directly to only the keypress task. Because vocal response latencies are believed to be somewhat slower than keypress latencies for the same class of task (e.g., Sipos, 1982), we added 50 ms to the value of the vocal motor component estimate when the vocal task was completed first. That is, for dual-task trials on which the keypress task was completed first, 200 ms was subtracted from the combined single-task sum and μ_{prep} estimate. Comparatively, for dual-task trials on which the vocal task was completed first, 250 ms was subtracted from the combined single-task sum and μ_{prep} estimate. Thus, the RT2 estimate based on the sum of the single tasks was corrected by subtracting a net value of only 100 ms (i.e., the estimate of $\mu_{\text{prep}} - [\mu_p + \mu_{km}]$ was -100 ms) when the keypress task was completed first on the corresponding dual-task trials, whereas the estimate was corrected by a net value of only 150 ms when the vocal task was completed first (i.e., the estimate of $\mu_{\text{prep}} - [\mu_p + \mu_{vm}]$ was -150 ms). As will become apparent later, these correction values are small relative to the observed effect sizes. As with the race model, the prediction for RT2 holds even if participants adopt a grouping strategy.

In addition to analysis of the mean, three other sources of evidence can be used to discriminate between the two classes of models. First, a race model must predict that the faster of the two single tasks will be completed first on the majority of dual-task trials. Any participant who completes the slower, vocal task first on most trials presents a serious problem for that model. Second, as discussed in detail later, the race and sequential models make divergent and testable predictions with regard to the correlation between RT1 and IRI (see also Rickard & Pashler, 2003). Third, we conducted RT distribution analyses to evaluate whether the conclusions reached on the basis of the means applied to the entire distribution.

It is important to emphasize that the predictions of both models assume that single-task associations are represented independently. This assumption is virtually guaranteed on the first dual-task trial for each item, because, before that point, the component single tasks have been performed independently and in different practice phases. The assumption may continue to hold beyond the first dual-task block. However, after the first dual-task block, the component tasks of the dual-task items will have been performed together previously. Learning on these previous dual-task trials could result in violation of the independence assumption, as for example in some type of response chunking effect. The strongest test of both models therefore involves the first dual-task block. This special case of component task independence is theoretically important. Most potentially applicable parallel models in the literature (e.g., Logan, 1988; Palmeri, 1997; Wenger, 1999) assume that parallel retrieval can take place in this case. Furthermore, a theoretical understanding of independent dual-task retrieval is a prerequisite to development of a model of dual retrieval practice effects, such as those observed in Experiment 2.

An implicit assumption of the estimation of dual-task RT predictions based on single-task performance is that participants use the same stopping rule for single- and dual-task blocks (Ashby & Townsend, 1986; Colonius, 1990; Van Zandt & Townsend, 1993). It is possible that this assumption is false. For example, participants could adopt a stricter stopping rule for the dual task, requiring more information accrual before executing responses than for single-task trials. In principle, the result could be dual-task RTs that are slower than even the sequential prediction outlined above, even though performance reflects parallel retrieval. Fortunately, this possibility can be tested. If participants did change their stopping rule in the manner outlined, it would probably be observable as a speed–accuracy trade-off such that error rates are lower on dual tasks than would be expected on the basis of single-task error rates. Beyond this, we must make the general assumption that the memory retrievals are processed in the same manner when they are performed on a dual-task trial as when they are performed on separate trials (e.g., Ashby & Townsend, 1986).

The RT1 and RT2 predictions of both models should be seen as boundary condition predictions. Each prediction specifies the lower bound mean dual-task RT that can be accommodated by the model. If the dual-task RTs fall significantly below one or more of these boundaries, then the corresponding model(s) can be rejected. Note that the sequential model should be easier to falsify on these grounds, because its boundaries are higher than those of the race model. On the other hand, dual-task RTs that are somewhat above a model boundary are not necessarily problematic for the principles underlying a model, because there is no a priori reason to expect that participants will perform at their maximum possible efficiency on all trials. The RT distribution fits should be quite useful in exploring this possibility. The possibility that participants might perform inefficiently on some dual task trials suggests that, if a model is correct, it should fit the data better on the lower regions of the distribution than on the upper regions.

Experiment 1

Experiment 1 addressed the question of whether extended single-task practice will allow for parallel dual-task performance. Participants practiced each of the two individual memory retrieval tasks until they began to reach asymptotic performance and then were given interleaved single- and dual-task blocks, as in the Rickard and Pashler (2003) study.

Method

Participants. Twelve University of California, San Diego undergraduates participated for course credit.

Materials and design. Each participant learned a vocal and a keypress response for each of 10 cues, as shown in Table 1. Each cue required a unique vocal response and one of two keypress responses. Cues were presented in white centered on a dark background using MEL (Psychological Software Tools, Pittsburgh, PA). A sample stimulus is shown in Figure 1.

RTs were measured with a voice key and a button box for the vocal and keypress tasks, respectively. An experimenter was present throughout the experiment, entering vocal responses and disqualifying from analysis those trials in which the voice key tripped too early or too late as a result of extraneous noise.

Learning and practice were divided over three 1-hr sessions. The first session consisted of a study phase and a practice phase, the second session consisted of extended practice, and the third session consisted of additional practice followed by the test phase. This design resulted in 40 single-task practice blocks followed by five triads of test blocks in which each triad contained 1 vocal task block, 1 keypress task block, and 1 dual-task block.

Table 1
Cue–Response Pairs for All Items: Experiments 1 and 2

Counterbalance Conditions A and B			Counterbalance Conditions C and D		
Cue	Vocal response	Keypress	Cue	Vocal response	Keypress
Red	5	←	Red	8	←
Green	4	←	Green	5	→
Blue	1	→	Blue	6	→
Yellow	3	→	Yellow	4	→
Purple	2	←	Purple	0	→
Brown	6	→	Brown	9	←
Black	7	←	Black	2	←
Orange	8	←	Orange	1	←
White	9	→	White	3	←
Pink	0	→	Pink	7	→

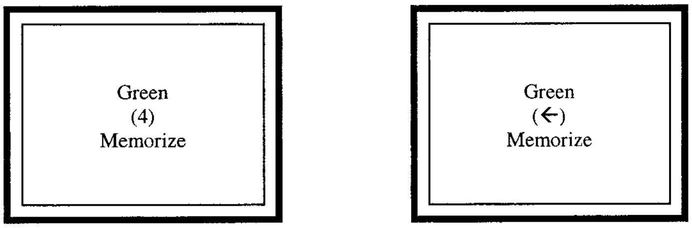
Procedure. After providing informed consent, participants were told they would be performing tasks involving learning new skills. To begin Session 1, participants were presented with each of the cue–response pairs of the vocal task in random order for 5 s apiece and were instructed to memorize them for later recall. After 5 s of study, the participant gave the appropriate vocal response by speaking the number into the voice key. This procedure was repeated for two blocks, with each block consisting of one instance of each of the 10 items, randomly ordered. After completion of these two study blocks, participants began the practice phase, in which they were required to generate the response. They were instructed to speak the correct number into the voice key as quickly and accurately as possible. Accuracy feedback, along with the correct response, was provided if participants made an error (as outlined in Figure 1). Participants received 10 blocks of vocal task practice. Next, the procedure just outlined was repeated for the keypress task.

To begin the second session, participants were told that the experimenters were interested in how fast people could perform the single tasks with practice. Participants practiced each of the tasks for 20 blocks, with alternating 10-block series for each task. They were given a brief rest after each 10-block sequence. At the beginning of the third session, participants again practiced 10 blocks of each task. This additional practice ensured that participants were performing at a nearly optimal rate when they entered the test phase. In the test phase, triplets of single- and dual-task blocks were interleaved (i.e., vocal, keypress, dual, etc.) a total of five times, with a final pair of vocal and keypress blocks added at the end. On the dual-task blocks, participants were instructed to both speak the correct number and press the correct button for each cue. Otherwise, instructions were the same for single- and dual-task blocks. Finally, we debriefed each participant.

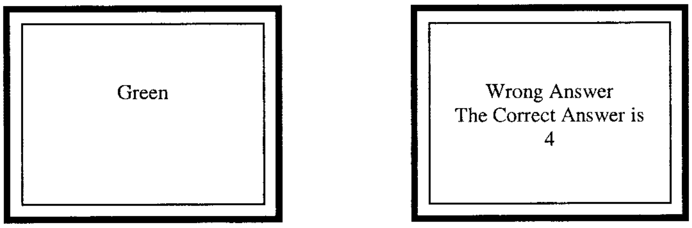
There were four counterbalancing conditions, each involving 3 participants. Participants in Conditions A and C completed tasks in the exact order just outlined. Participants in Conditions B and D completed the vocal and keypress tasks in reverse order. In Conditions C and D, the stimulus–response pairings were randomly changed relative to those of Conditions A and B.

Results and Discussion

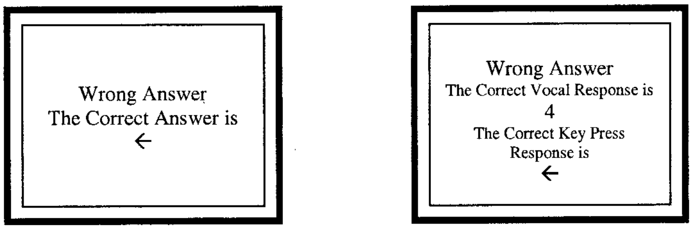
The voice key failed on only 0.2% of trials. During the course of practice, mean correct RTs in each block decreased from 1,061 ms to 681 ms for the vocal task and from 638 ms to 490 ms for the keypress task. The speed-up functions for each single task are shown in Figure 2. Accuracy by the end of practice was near 100%. In an informal interview after the experiment had been



“Memorize” Stimulus – Vocal “Memorize” Stimulus – Key Press



Test Stimulus – All Tasks Error Feedback – Vocal Task



Error Feedback – Key Press Error Feedback – Dual Task (2 errors)

Figure 1. Sample stimuli.

completed, many participants reported frequent use of mnemonics to establish the new cue–response association during initial single-task practice blocks. For example, one participant reported retrieving “golf” when presented with the cue “green,” which in turn cued “fore,” leading to the vocal response “four.” By the conclusion of practice, however, participants who initially used mnemonic strategies almost unanimously reported that they had discarded them and were retrieving the answers directly (i.e., without intermediate thought processes). This finding, combined with the fast and stable RTs at the end of single-task practice, strongly suggests that single-task retrieval at test was automatic by most any current definition (e.g., Logan, 1988).

During the test phase, 96% of all trials were performed correctly. The vocal task alone was performed at 98% accuracy, and the keypress task at 96%. In contrast, both responses were correct in the dual task on only 92% of trials. Performance on the dual task was only slightly less accurate than predicted by the product of single-task accuracies (94%). This accuracy level is predicted for the dual task because both models assume that component task retrievals are independent. A within-subjects analysis of variance (ANOVA) with block and data type (dual-task accuracy vs. predicted accuracy) as factors was performed on the proportion of correct responses computed for each participant on each test block. There was no significant effect of data type, $F(1, 11) = 1.27$, $p = .28$, nor was there an interaction with block, $F(4, 44) = 1.22$, $p = .31$. As noted earlier, these findings speak against the possibility of any serious violation of context independence in the single- versus dual-task comparisons.

Overall, participants executed the vocal response first on 32% of dual-task trials. However, most participants consistently performed the dual task using either one response sequence or the other. Three of the 12 participants executed the vocal response first on more than 88% of trials, whereas 6 others executed the keypress task first on more than 92% of trials.

Figure 3 shows RTs for both the single and dual tasks in the test phase. The dual-task RTs, along with the sequential and parallel model predictions for RT1 and RT2, are shown in Figure 4. As noted earlier, performance on each dual-task block for each item was predicted on the basis of performance on each of the two immediately preceding single-task blocks for that item (one vocal and one keypress block). This approach allowed us to match adjacent single- and dual-task triplets and analyze only those for which all three members were responded to correctly. Thus, the

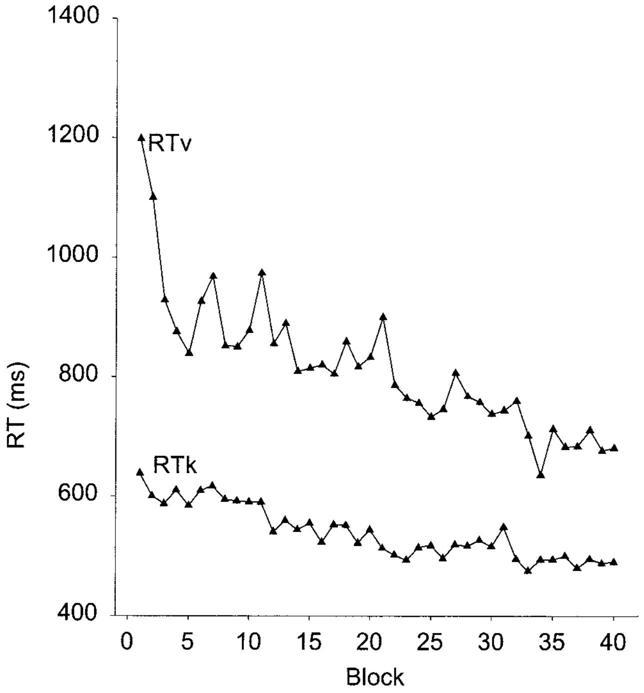


Figure 2. Single-task response times (RTs) in the practice phase of Experiment 1. RTv = vocal trial; RTk = keypress trial.

results described subsequently were not confounded by missing values.

We performed two 2-way factorial ANOVAs for both RT1 and RT2, one comparing mean RT values with the predictions of the sequential model and the other comparing the RT values with the predictions of the race model. The factors were block (1 to 5), data type (observed vs. predicted), and their interaction. In the analysis comparing RT2 with the sequential prediction, there was no significant effect of data type, $F(1, 11) = 0.03$, $p = .87$, but there was a significant interaction between data type and block, $F(4, 44) = 7.94$, $p < .0001$. In the analysis comparing RT2 with the parallel prediction, there were significant effects for both data type, $F(1, 11) = 9.59$, $p = .01$, and the interaction of data type with block, $F(4, 44) = 9.20$, $p < .0001$. RT1 was significantly different from the sequential prediction, $F(1, 11) = 14.59$, $p = .0029$, and there was a significant interaction with block, $F(4, 44) = 4.96$, $p = .0022$. Finally, RT1 deviated strongly from a parallel model, $F(1, 11) = 18.44$, $p = .0013$, and a significant interaction with block was again present, $F(4, 44) = 5.21$, $p = .0016$. All of these effects reflect the patterns apparent in Figure 4.

To evaluate the range of empirical predictions of possible limited capacity parallel models that can be ruled out, we calculated the difference between RT2 and the sequential model's prediction for each participant's first test block. A 95% confidence interval for this difference value was then computed. The interval had a lower bound of -103 ms, corresponding to a lower bound for RT2 of 1,045 ms. In contrast, the race prediction for these same data was 858 ms. This result indicates that we can reject not only the special case of a race model but also any limited capacity parallel model that would yield a prediction for RT2 between 858 ms and 1,045 ms on the first test block.

The RT analyses just discussed are clearly more consistent with the sequential model than the race model. On the first dual-task block, on which the tasks were presumably independent, RT1 and RT2 were both above the lower bound prediction of the sequential model (indicated as sequential model RT1 and RT2 predictions in Figure 4). On subsequent blocks, RT2 fell below that prediction to a modest extent.

It is, of course, possible that the results in regard to the means do not hold up across the entire distribution. In particular, it is possible that, at the lower tail of the distribution, RT1 or RT2 (or both) is faster than the sequential prediction. A significant crossover on the lower tail on the first block would falsify the sequential model and might be indicative of some type of inefficient parallel retrieval process. On the other hand, if RT2 is slower than the sequential prediction at all points on the distribution, then our conclusion in favor of the sequential model would be substantially reinforced. We plotted cumulative distributions for both RT1 and RT2 on Block 1, alongside their respective parallel and sequential predictions, following the procedures of Ratcliff (1979) and Miller (1982). The 10 RTs (separately for RT1, RT2, and all four of the predictions) for each participant were rank ordered from fastest to slowest. Inclusion of all responses, regardless of accuracy, was necessary to maintain a valid RT ordering for each participant. Incorrect responses were distributed roughly equally throughout all 10 decile values along the distribution, so their inclusion would not have been expected to bias the results in any particular direction.

The rank-ordered data were then averaged across participants, yielding one group mean for each of the 10 deciles. Recent work

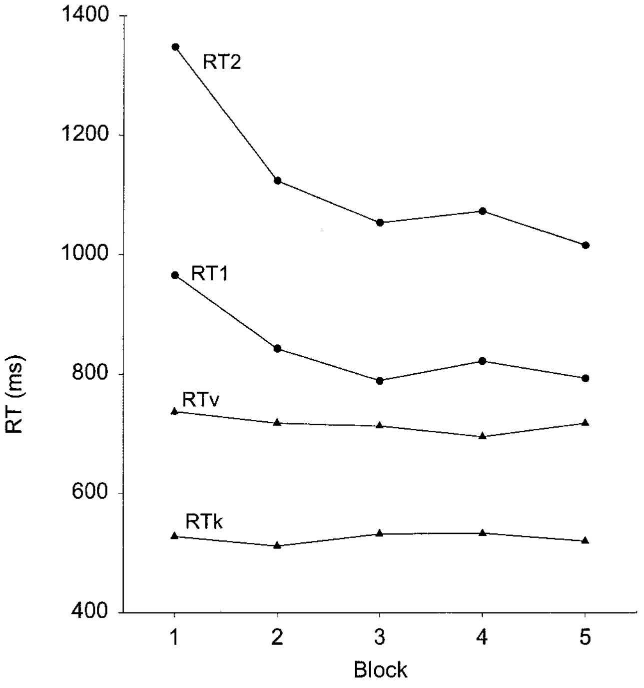


Figure 3. Single-task and dual-task response times (RTs) in the test phase of Experiment 1. RT1 = latency to first response on a dual-task trial; RT2 = latency to the second response on a dual-task trial; RTv = vocal trial; RTk = keypress trial.

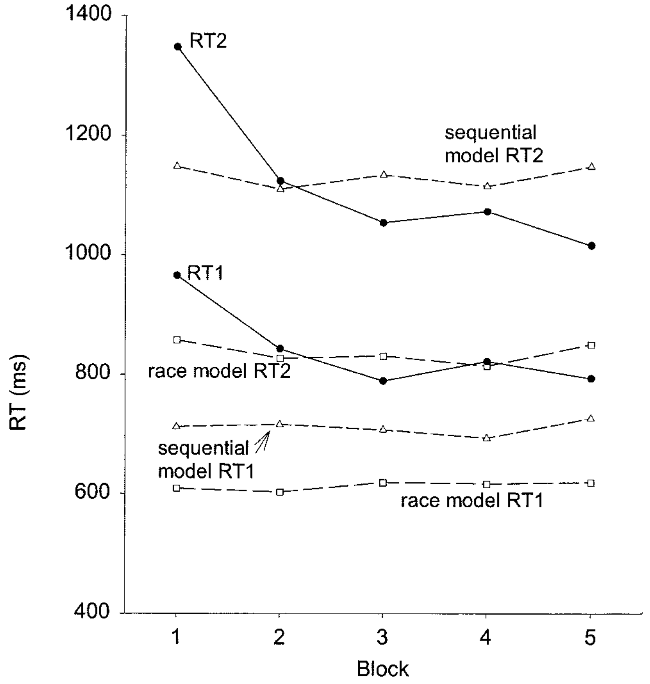


Figure 4. Comparisons between Experiment 1 response times (RTs) and predictions. RT1 = latency to first response on a dual-task trial; RT2 = latency to second response on a dual-task trial.

has shown that attempts to either reproduce the shape of the population distribution or estimate the parameters of a candidate distribution function can lead to systematic biases, especially with small samples (Van Zandt, 2000). However, in this case neither was attempted. Rather, the goal was the much simpler one of determining whether or not the critical RT effects at the level of the mean held for the entire distribution. If there is no cumulative distribution crossover in the populations, then no crossover should be present in the sample, and thus none should be observed in the plots (Townsend, 1990).

Figure 5 shows the first block cumulative distribution for RT1, along with the predictions of both models. At no point on the distribution does RT1 fall below the sequential prediction. In fact, RT1 is substantially greater than both the sequential and parallel predictions throughout the distribution, mirroring the pattern for the means. One possible explanation of this finding is that some participants adopted a grouping strategy, resulting in an artificial delay of RT1 (see also Pashler & Johnston, 1989). Individual-level IRIs did suggest that some participants adopted a grouping strategy, slowing RT1, although RT2 was unaffected by this distinction.

The RT2 values, shown in Figure 6, were greater than the predictions of both models on the upper tail of the distribution but converged onto the sequential prediction toward the lower tail. A matched *t* test was used to compare each observed decile value with the sequential prediction for that decile (see Miller, 1982, for precedent and for simulation results confirming containment of Type I error rates when this procedure is used). At the lowest decile, RT2 was not significantly different from the sequential prediction, $t(11) = 0.34$, $p = .74$, nor was the difference significant for any decile up to and including the seventh. This pattern suggests that the low end of the distribution represents highly efficient processing of one retrieval, followed by the other retrieval in sequence, with no excess RT component due to factors such as hesitation or dual-task confusion (for a related discussion, see Schweickert, 1983). Note that, from the standpoint of a limited capacity model, the sequential boundary onto which RT2 converges is arbitrary. If performance were to reflect limited capacity retrieval, this convergence would not be expected. It appears instead that the sequential model provides a theoretically meaningful lower bound on performance.

An additional piece of evidence favoring the sequential model comes from the patterns of correlation between RT1 and IRI on the dual-task trials. As noted previously by Rohrer et al. (1998) and Rickard and Pashler (2003), the race model predicts a negative correlation between RT1 and IRI. If the two retrievals "race" independently, and if RT1 on a particular trial is unusually slow, then on average the IRI will be unusually fast (because on average RT2 would not be expected to be especially slow on that trial). The reverse is also true. Rohrer et al. (1998) showed, through simulations, that the expected magnitude of the correlation is about −.5. The sequential processing model also assumes that the two retrievals are independent, but crucially, of course, it assumes that they are performed sequentially. For participants who always complete one of the tasks (vocal or keypress) first (see Schweickert, 1983, for a related case in which task ordering is crucial), the sequential processing model predicts zero correlation between RT1 and IRI.

On the other hand, because the means for these two tasks are different, participants who mix their response ordering (i.e., some with keypress first, and some with vocal first) will show a negative

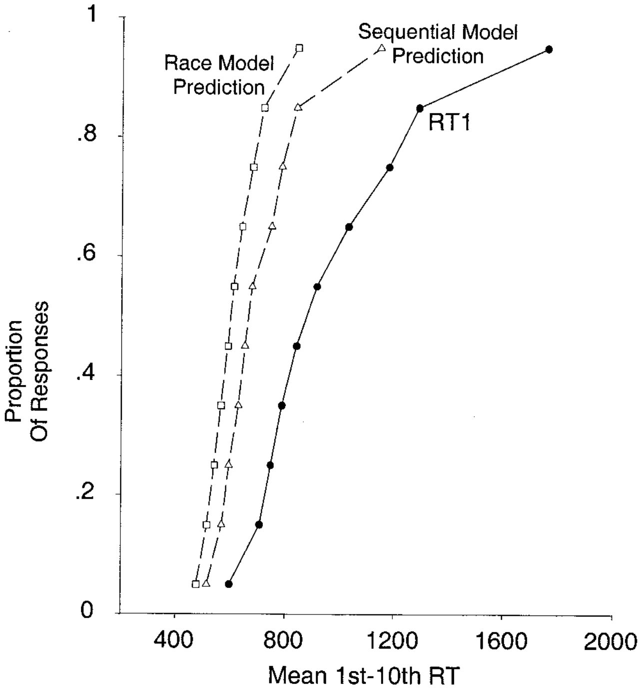


Figure 5. Cumulative distribution of RT1 on Block 1 of the test phase of Experiment 1. RT = response time; RT1 = latency to first response on a dual-task trial.

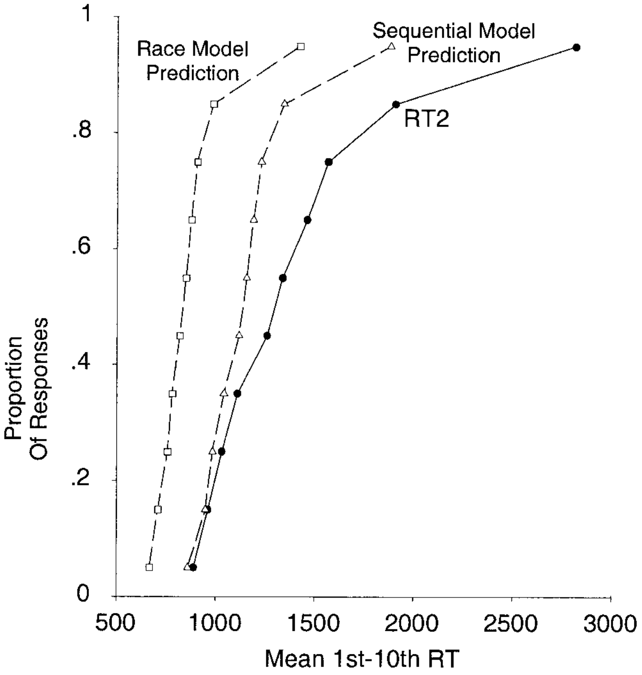


Figure 6. Cumulative distribution of RT2 on Block 1 of the test phase of Experiment 1. RT = response time; RT2 = latency to second response on a dual-task trial.

correlation even if the sequential model is correct. We thus limited this analysis to participants who always completed the same task first, be it vocal or key. Only 2 of the 12 participants qualified for the correlation analysis by this criterion. We analyzed the dual-task data separately for each of these participants, treating IRI as the dependent variable, block and item as categorical covariates of no interest (to factor out practice and item difficulty confounds), and RT1 as a continuous independent variable of interest. The partial correlations between RT1 and IRI were nonsignificant for both participants (*p*s = .77 and .41), converging on the results from multiple participants in the Rickard and Pashler (2003) study.

As noted earlier, 3 participants executed the vocal task response (the slower of the two tasks on average) first on nearly all of the dual-task trials. No parallel model, not even of a limited capacity variety, can account for this finding unless a strategic capacity allocation mechanism is included. Such a mechanism would allow participants to allocate most of their attention to the slower task. However, that model would be more complex and would appear to be quite flexible (see Rickard & Pashler, 2002, for a more detailed critique of limited capacity parallel models as applied to this task).

Experiment 2

The remarkable magnitude and resilience of dual retrieval slowing observed in both the Rickard and Pashler (2003) experiments and the preceding experiment raises the question of whether there are any conditions under which participants will be able to perform two retrievals in parallel, with corresponding RT facilitation, when given a single cue. In Experiment 2, we explored this question for the case of extended dual-task practice. Several possible outcomes seem viable on the basis of the results observed so far. First, it is conceivable that performance will continue to be more consistent with the sequential than with the race or other models even after extensive dual retrieval practice. Such a result would suggest a very robust retrieval bottleneck that is probably operative under most conditions. A second possibility is that independent parallel retrieval can take place after extensive dual-task practice, allowing dual-task RTs to approach the predictions of the race model. Alternatively, some type of interactive, or chunked, representation may develop with practice, resulting in dual-task RT facilitation. Finally, it is possible that dual-task performance will depend crucially on whether participants adopt a grouped or a nongrouped response strategy.

Method

Participants. Sixteen University of California, San Diego undergraduates participated for course credit.

Materials, design, and procedure. The materials were identical to those used in Experiment 1. The design and procedure were also similar to those of Experiment 1, with a few exceptions. First, the practice and test phases were the same in this experiment; participants began interleaved single- and dual-task practice soon after the study blocks. We refer to this phase as the test. Second, after the study phase and before the test phase, participants were presented with a few blocks of single-task items for each task, as in the practice phase of Experiment 1. In this case, however, the goal was simply to ensure that participants reached a minimum proficiency on the single-task items before proceeding to the test phase. During this phase, the participants and the experimenter were given accuracy and mean RT feedback at the end of each block (there was also trial-based error feedback as in Experiment 1). A task was considered to have been learned when the participant performed two consecutive blocks with 100% accuracy, the second with a mean RT below 1,200 ms. These same criteria were used for the vocal task and the keypress task. In the following test phase of Session 1, participants received 5 triads of each task (as in the test phase of Experiment 1). The second session consisted of an additional 15 test triads, with a short break after every 5 triads. Finally, the third session consisted of 10 more triads, again with a break after 5 triads. This procedure resulted in a total of 30 test triads over three sessions.

Results and Discussion

The mean numbers of blocks needed to reach the single-task performance criteria were 4.9 in the vocal task and 3.4 in the keypress task. All of the results described subsequently involve the test phase. The voice key failed on 0.2% of responses, and these data were excluded from analyses. The overall accuracy rate was 98%, with accuracy on the vocal task alone at 98% and overall keypress accuracy at 97%. Both responses were correct on 97% of dual-task trials. Although dual-task accuracy was slightly higher than that predicted under an assumption of independent retrieval (95%), an ANOVA (as described for the error analysis of Experiment 1) revealed no significant effect of either data type (observed accuracy vs. predicted accuracy), $F(1, 15) = 0.70$, $p = .41$, or the interaction of data type with block, $F(29, 435) = 1.17$, $p = .25$.

The vocal response was executed first on 13% of all dual-task trials. This figure varied for individual participants. The majority of participants (11 of 16) executed the keypress first on more than 95% of dual-task trials, whereas 1 participant executed the vocal response first on more than 99% of dual-task trials.[3] Data trios with one or more incorrect responses were excluded from mean RT analyses according to the procedure described in Experiment 1. RTs for each single task and for the dual task are plotted as a function of practice block in Figure 7. Over the course of practice, mean RTs decreased from 1,049 ms to 677 ms on the vocal task and from 834 ms to 512 ms on the keypress task. RT1 means decreased from 1,483 ms to 606 ms, and RT2 means decreased from 2,046 ms to 809 ms.

Figure 8 compares RT1 and RT2 with the parallel and sequential model predictions. As in Experiment 1, several two-way ANOVAs were performed, comparing observed RT2 and RT1 with sequential and parallel model predictions, again with data type (observed vs. predicted) and block as variables. In the comparison of RT2 with the sequential prediction, both the main effect of data type, *F*(1, 15) = 13.94, $p = .002$, and the interaction of data type with block, *F*(29, 435) = 5.45, $p < .0001$, were significant. In a second analysis comparing RT2 with the parallel prediction, the main effect, *F*(1, 15) = 20.67, $p = .0004$, and the interaction, *F*(29, 435) = 11.16, $p < .0001$, were both strongly significant. A third analysis compared RT1 with the sequential prediction. There were again significant effects of both data type, *F*(1, 15) = 13.34, $p = .0024$, and the interaction with block, *F*(29, 435) = 10.56, $p < .0001$. Finally, an analysis comparing RT1 with the parallel prediction again revealed a significant main effect of data type, *F*(1, 15) = 17.11, $p = .0009$, and an interaction, *F*(29, 435) = 14.70, $p < .0001$. These results quantitatively confirm the major patterns visible in Figure 8. In short, neither model fits the overall data set particularly well.

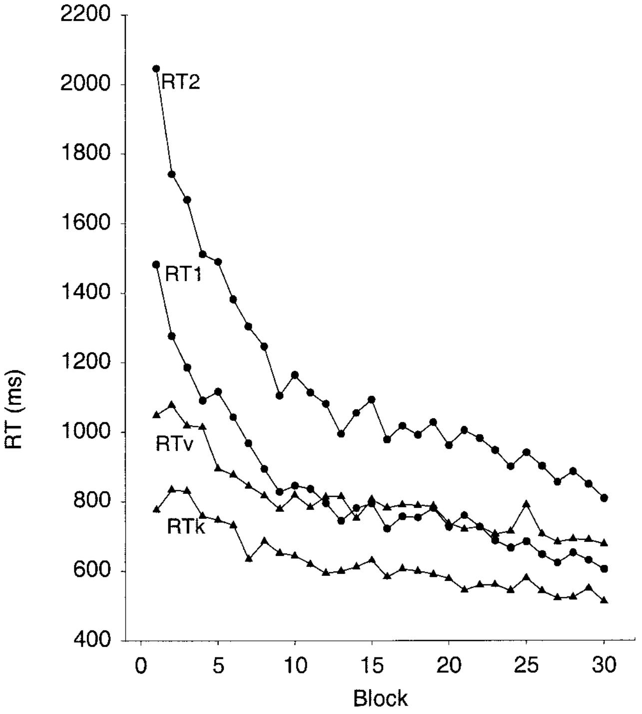


Figure 7. Single- and dual-task response times (RTs) in Experiment 2. RT1 = latency to first response on a dual-task trial; RT2 = latency to second response on a dual-task trial; RTv = vocal trial; RTk = keypress trial.

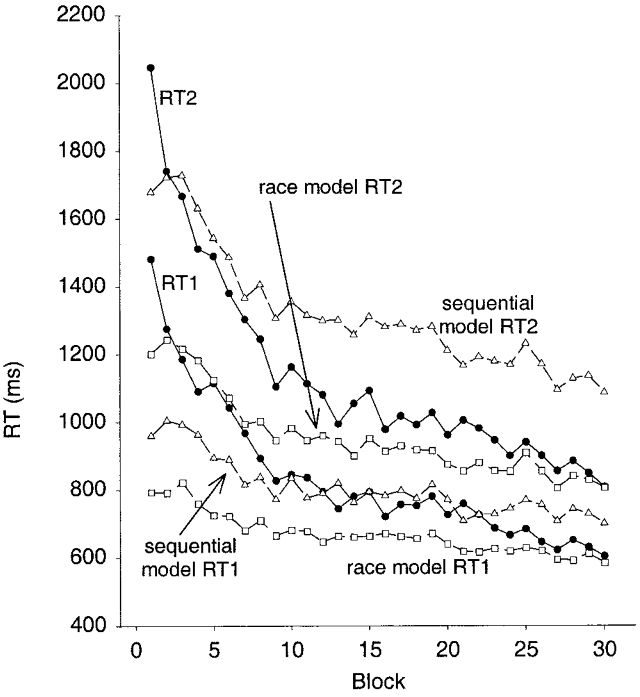


Figure 8. Comparisons between Experiment 2 response times (RTs) and predictions. RT1 = latency to first response on a dual-task trial; RT2 = latency to second response on a dual-task trial.

However, as with Experiment 1, the first block was the only one on which the two responses were reasonably certain to be independent, and thus the first block provided the best test of the models. The confidence interval on this block of the difference between RT2 and the sequential prediction (calculated as in Experiment 1) had a lower bound of 112 ms. In other words, the lower bound of the 95% confidence interval was 1,791 ms, and this value was above the sequential prediction, 1,679 ms. Thus, any limited capacity parallel model that predicts subsequential dual-task performance can be rejected. The sequential model in its simplest form can also be rejected for mean RT2 on the first block. However, it remains possible that, as in Experiment 1, these unusually slow RTs reflect inefficiencies in processing on some dual-task trials.

To further investigate dual-task performance on the first block, we plotted cumulative distributions following the procedure of Experiment 1. The result for RT1 is shown in Figure 9, and the result for RT2 is shown in Figure 10. The patterns strongly resemble those seen in Experiment 1. Figure 9 shows a delay in RT1 that may be attributable to some of the participants' grouping responses. Thus, it is difficult to make theoretical inferences from

---

[3] Three of the remaining 4 participants changed their response ordering as the experiment progressed. During the first five blocks, these 3 participants executed the vocal response first on 59% of the dual-task trials. On the next five blocks, this percentage dropped to 36%. On Blocks 11–20 the vocal task was faster on only 13% of dual-task trials, and on the last 10 blocks these 3 participants gave the vocal response first on only 8% of dual-task trials.

Figure 9. As shown in Figure 10, RT2 converges onto the sequential prediction at the low end of the distribution. At neither of the lowest two deciles was RT2 significantly different from the sequential prediction, $t(15) = 1.03$, $p = .31$, and $t(15) = 2.07$, $p = .055$, for the first and second deciles, respectively. As in Experiment 1, it appears that when participants were processing the component tasks most efficiently, they were able to perform the dual task in a nearly optimal sequential fashion. A priori, these results would not be expected by a limited capacity parallel retrieval account.

The possibility that some of the participants adopted a grouping strategy, as suggested by Figures 8 and 9, was investigated. Following Rickard and Pashler (2003), each participant was categorized into one of two groups according to mean IRI. Participants with mean IRIs of less than 300 ms were classified as groupers and were assumed to have grouped their response outputs on most if not all trials, delaying RT1 output until both responses were retrieved and then executing both motor responses nearly simultaneously. Ten of the 16 participants fell into this category. The remaining 6 participants, who had mean IRIs of more than 300 ms, were classified as nongroupers. In support of this classification boundary, the mean IRIs were bimodally distributed, with no mean IRIs in the range between 218 ms and 436 ms.[4] The grouping classification did not appear to be related to response ordering. The 6 nongroupers executed the vocal response first on 19% of dual-task trials, and the 10 groupers executed the vocal response first 9% of the time.

Separate results for nongroupers and groupers are presented in Figures 11 and 12, respectively. Examination of these graphs suggests that the failure of the overall mean RTs to match either model's predictions was largely the result of averaging over the different response output strategies. Among nongroupers, RT2 was similar to the sequential prediction throughout most of practice. In the case of these participants, there was no main effect of data type (observed vs. predicted), $F(1, 5) = 0.02$, $p = .88$. However, there was an interaction of data type with block, $F(29, 145) = 3.90$, $p < .0001$. Among nongroupers, RT1 also did not significantly differ from the sequential prediction, $F(1, 5) = 1.45$, $p = .28$, although again an interaction with block was present, $F(29, 145) = 3.24$, $p < .0001$. The similarity between nongroupers' observed RT2 and the sequential prediction is visible in Figure 11.

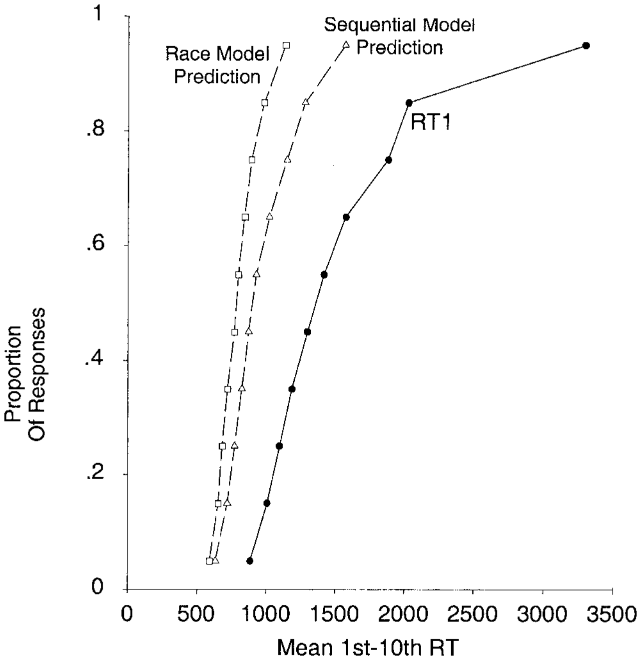


Figure 9. Cumulative distribution of RT1 on Block 1 of Experiment 2. RT = response time; RT1 = latency to first response on a dual-task trial.

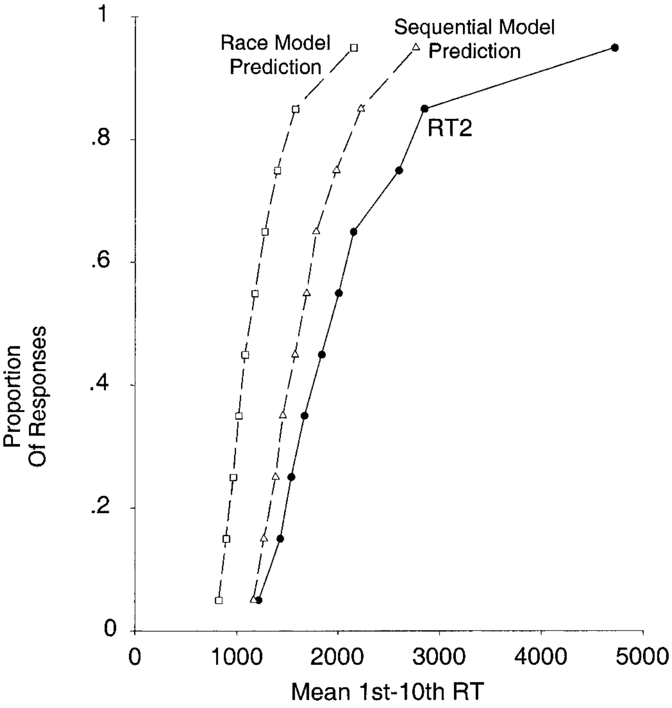


Figure 10. Cumulative distribution of RT2 on Block 1 of Experiment 2. RT = response time; RT2 = latency to second response on a dual-task trial.

Despite the interaction with block, the sequential model undoubtedly provides a better fit to the data for nongroupers than does the race model. Nevertheless, dual-task performance fell below both the RT2 and RT1 predictions of that model by the end of practice. One plausible interpretation is that retrieval is no longer sequential at the end of practice, even for nongroupers. Although this possibility cannot be ruled out, inspection of IRIs as a function of practice, as compared with the IRI values predicted by both models (Figure 13), provides intriguing evidence to the contrary. With the exception of the first five or six blocks of practice, the sequential prediction is quite accurate. Note that the model provides this fit without the benefit of free parameters. Further, this analysis cannot be confounded by any possible error in our estimation of preparation cost, and is immune to any

---

[4] Mean IRIs were 39, 97, 102, 114, 131, 153, 164, 164, 186, and 218 ms for the 10 groupers, and 436, 485, 504, 508, 662, and 785 ms for the 6 nongroupers.

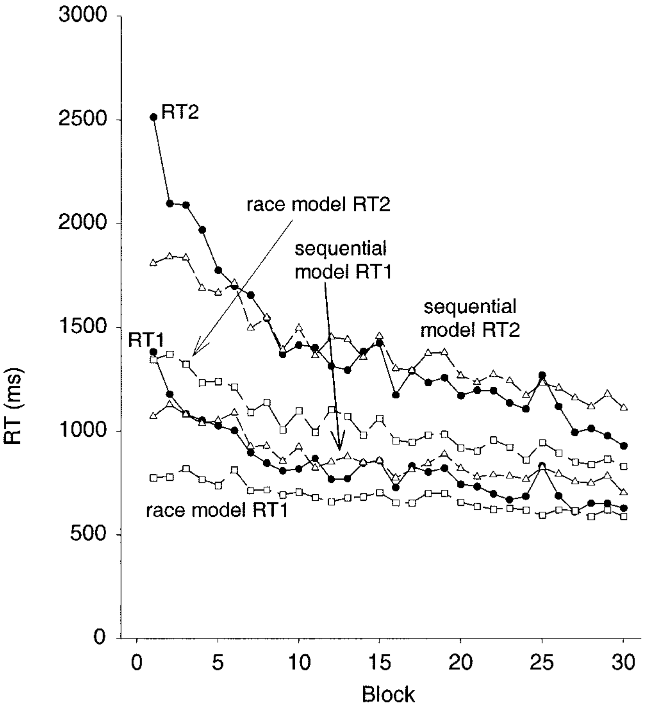


Figure 11. Comparisons between nongroupers' response times (RTs) and predictions: Experiment 2. RT1 = latency to first response on a dual-task trial; RT2 = latency to second response made on a dual-task trial.

inefficiencies in processing that occur prior to completion of the first task. As such, the results shown in Figure 13 provide solid converging evidence in favor of sequential retrieval.

Although the IRI results support the sequential account, the deviations of RT1 and RT2 from that model's predictions at the beginning and end of practice, even for nongroupers (Figure 11), remain to be explained. We proposed earlier that the model's underprediction of RT2 and, to a lesser extent, RT1 at the beginning of practice is the result of inefficiencies in management of the dual task, as suggested for Experiment 1. This conclusion is consistent with the distribution fits for RT2 in both experiments (Figures 6 and 10). On the lower tail, RT2 converged with the sequential prediction, suggesting that participants were sometimes capable of optimal sequential retrieval even on the first dual-task block.

With modest additional practice, both RT1 and RT2 for nongroupers converged on the sequential prediction, as shown in Figure 11, suggesting that these participants learned to perform the dual task efficiently a greater proportion of the time. This close fit for RT2 held until about Block 15. Beyond that block, a developing trend for the sequential model to overpredict both RT1 and RT2 was evident. By the end of practice, the average amount of overprediction was about 100 ms. One interesting possibility is that these RT decreases reflect a gradual dissipation of the dual-task preparation effect with practice, similar to the dissipation in preparatory effects in other tasks demonstrated by Gottsdanker (1975) and by Van Selst, Ruthruff, and Johnston (1999). The IRIs plotted in Figure 13 support this interpretation. That analysis eliminated the need to estimate the preparatory cost, and the data fit the sequential model throughout almost all of practice.

As in Experiment 1, we searched for correlations between RT1 and IRI in the dual-task data in an effort to test a limited capacity account. This analysis was performed for each nongrouper who always, or nearly always, completed one of the tasks (vocal or keypress) first across all test blocks. The regression model was identical to that used in Experiment 1. The effect of RT1 on the IRI was not significant for any of the 5 participants who met the criteria (p values were .06, .14, .15, .25, and .52). This result converges with that of Experiment 1, as well as with the correlation results of Rickard and Pashler (2003).

It is interesting that both RT1 and RT2 for the groupers (Figure 12) roughly matched the parallel prediction during the last half of practice. It may be that, after considerable practice, groupers were able to retrieve both answers independently and in parallel, just as the race model predicts. It should be noted that if retrieval is parallel, and if there is a positive correlation between the keypress and vocal task retrieval times on dual-task blocks, the resulting RT2s will be somewhat below the race model prediction, and the resulting RT1s will be somewhat above it (Colonius, 1990), as shown at the end of practice in Figure 12. This possibility is intriguing and merits further investigation. On the other hand, the onset of chunked or interactive task processing with practice provides an equally viable hypothesis. According to this account, retrieval on dual-task trials is parallel, but the "channels" through which the two tasks are retrieved are no longer separate. Rather, there is cross communication of some sort between the tasks during the retrieval stage of processing (Ashby & Townsend, 1986; Colonius, 1990; Diederich & Colonius, 1987).

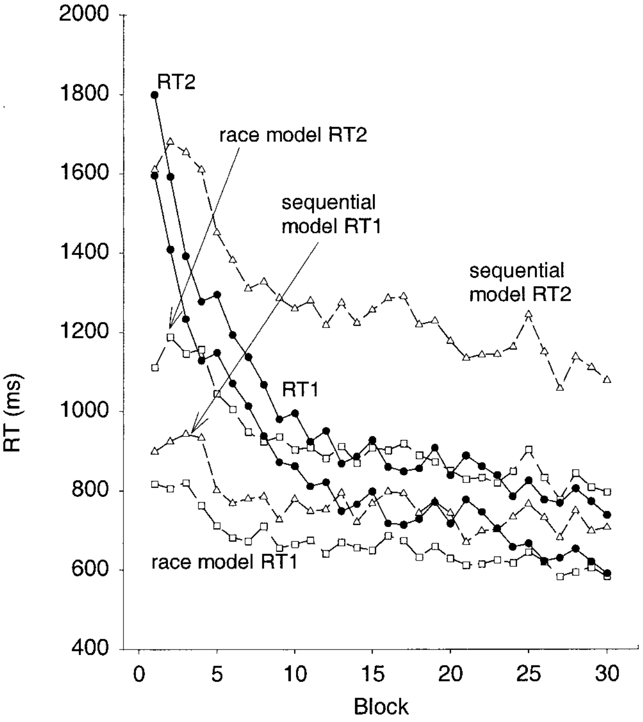


Figure 12. Comparisons between groupers' response times (RTs) and predictions. RT1 = latency to first response on a dual-task trial; RT2 = latency to second response on a dual-task trial.

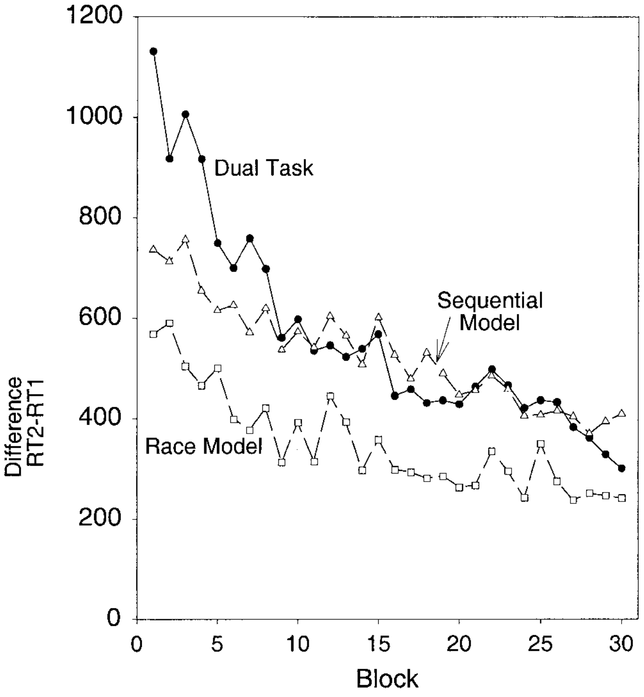


Figure 13. Difference scores for nongroupers (RT2 − RT1), predicted and observed: Experiment 2. RT1 = latency to first response on a dual-task trial; RT2 = latency to second response on a dual- task trial.

General Discussion

These experiments contribute three new findings to the literature on memory and attention while also uncovering an interesting new learning phenomenon that merits further investigation. First, it is now clear that the sequential retrieval observed by Rickard and Pashler (2003) is not limited to a special case of minimal single-task practice. In Experiment 1 of the present study, participants were given 40 blocks of single-task practice, resulting in nearly perfect accuracy as well as large reductions in both RT and subjective effort. Yet, their performance was consistent with the sequential model prediction on the first test block and closer to the sequential than to the parallel prediction for the remainder of the test. This finding has notable implications with respect to automaticity theory, suggesting that the basic processes underlying retrieval do not change qualitatively with single-task practice. More specifically, our results suggest that the retrieval bottleneck does not dissolve after practice on single tasks. The transition from mnemonically mediated to subjectively more direct retrieval, for example, does not appear to open the door to parallel memory retrieval, at least not for the current task. If one endeavors to treat memory retrieval as an automatic process, then our results appear to require that one property traditionally associated with automaticity, namely parallel processing, must be dropped, at least for the case of two retrievals from a single cue. Practically speaking, the results of Experiment 1 suggest that the limitation of retrieving through only one independent pathway at a time may be the general rule in everyday cognition.

A second strong conclusion is that even dual retrieval practice does not guarantee that participants will exhibit retrieval-related performance savings relative to performing the two retrievals on separate trials. This result is surprising and is a further testament to the robustness of the retrieval bottleneck. Third, the results imply that performance on the dual task, and in particular the skill acquisition effect with practice, is strongly correlated with the "strategy" that a participant adopts (whether intentionally or not). If and only if participants choose to group their response outputs can their dual-task performance (RT2) fall substantially below the sequential prediction, and even then sufficient dual-task practice is required.

The current results, combined with those of Rickard and Pashler (2003), provide considerable support for a sequential processing model of dual-task performance when task associations are independent, regardless of the amount of single-task practice. The results of other studies conducted in our laboratory show that these findings are not specific to the stimulus–response mapping used in the current study. Rickard and Pashler (2003, Experiment 2) obtained the same results in a task involving eight stimuli, eight verbal responses, and eight keypress responses (as opposed to two keypresses in this task). Other unpublished research from our laboratory shows that the results are robust to changes in instructions to participants. In one condition of an unpublished study, participants were told that, in previous research, it was found that individuals could perform the two retrievals at the same time if they intentionally attempted to do so. These participants were instructed to make this effort. The results for that condition were similar to those for participants in the neutral instructions condition of the experiment (the same instructions that we used in the current experiments). The performance of both groups was again consistent with the sequential model.

Nevertheless, the sequential model cannot account for the substantial dual-task performance speedup observed for groupers with dual-task practice in Experiment 2. Other unpublished work has demonstrated that participants who are instructed to group their responses show a pattern of reduction in RT2 very similar to that of groupers in Experiment 2, indicating that response strategy differences in the current study are not due to intrinsic individual differences in processing capacity or skill. It appears instead that the response strategy that participants choose to use is the causal factor underlying the corresponding differences in RTs.

We consider next two general approaches to account for the response grouping effect. One approach assumes a hybrid model in which a bottleneck is present at the beginning of practice, only to dissolve and allow for parallel retrieval through separate, independent channels at some point thereafter. It is of interest here that similar transitions from serial to parallel processing with practice appear to occur in visual search and, perhaps, other task domains (see Logan, 1978). However, this hybrid model poses two fundamentally different types of retrieval (at both cognitive and neural levels), and it leaves open the question of how the transition between these diametrically opposed types of processing might be implemented. It also raises the question of why independent parallel retrieval would occur only when participants adopt a grouping strategy *and* have had substantial dual-task practice.

Alternative accounts of our data preserve the idea of an immutable retrieval bottleneck, regardless of amount of single- or dual-task retrieval practice. As a preliminary, consider the hypothesis that both responses must be held concurrently in the focus of working memory before new learning, leading to subsequential performance, can take place. This condition is, of course, most likely to have been met for participants who grouped their responses. Given this condition, two hypotheses can potentially account for the overall patterns observed in the data without abandoning the retrieval bottleneck premise.

Both of these hypotheses can be conceptualized as extensions of the model of memory and skill learning proposed by Rickard (1997, 1999; see also Rickard & Bajic, in press), although other theoretical frameworks would probably suffice. Our goal here is not to validate any particular model but, rather, to provide concrete examples of models that are at least sufficient to generate the basic pattern of results. First consider Figure 14A, which shows the associative configuration expected by the model after single-task learning. At the input level, the model assumes separate representations for the cue and the goal, or task set (i.e., “retrieve vocal response” or “retrieve keypress response”). The model assumes that learning results in a “set-cue” level of representation (termed the “problem level” by Rickard, 1997) that corresponds to the conjunction of each stimulus and task goal in the figure. In turn, this set-cue node for each item is associated with the required response (for more detail, see Rickard, 1997, and Rickard & Bajic, in press).

Importantly, the model assumes that only one set-cue node at a time can mediate performance. Hence, the source of the retrieval bottleneck is located at the set-cue level in Figure 14A. This model requires that, to execute both retrievals in the case of task independence, the participant must cycle through this network twice, once for the keypress task and again for the vocal task. If one assumes that the system “resets” to the default state after completion of each retrieval, then this model predicts sequential retrieval and is consistent with Equations 5 and 6.

One way in which this model might be extended to provide an account of our dual-task practice results involves the added assumption that, in the special case in which both task sets and both responses are concurrently present in the focus of working memory (i.e., for participants using a grouping strategy), a new node is created that is associated with the cue and both task goals (see Figure 14B). With sufficient practice, there may be a transition from two sequential retrievals (e.g., moving first through the keypress pathway and then the vocal pathway) to a single retrieval mediated by the new dual-task node (for a discussion of similar strategy transitions, see Rickard, 1997, 1999). On any given trial, retrieval would take place either sequentially through the original, single-task set-cue nodes or at once through the new, dual-task set-cue node. Although the details remain to be worked out, this approach potentially allows for the eventual onset of subsequential retrieval without violating the retrieval bottleneck hypotheses. In essence, it reduces the dual task to a single retrieval of two responses, with only one set-cue node mediating performance on any trial. Dual retrieval of this sort would resemble independent parallel retrieval of the two responses, but the responses are in fact not being retrieved independently throughout the entire pathway from stimulus to response. Such a process could be described as parallel, but it would not involve two independent parallel retrievals in the same sense as in a race or as in a limited capacity parallel account in which the two retrievals are functionally independent.

Another possible extension of the model eschews the formation of a new node at the problem level in favor of an associative chaining account at the response level (see Figure 14C). Here, participants are initially retrieving sequentially on the dual task. Provided that they group their outputs, however, a direct association forms between the two responses, perhaps within some part of

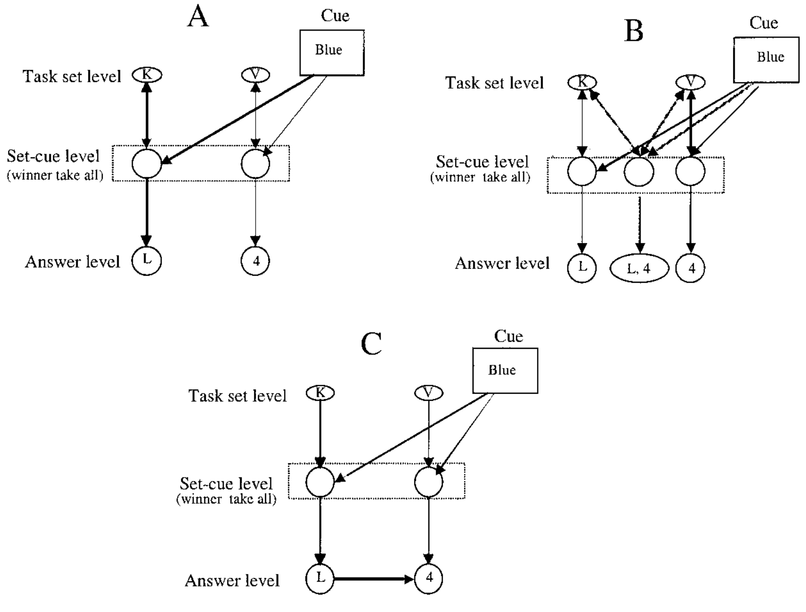


Figure 14. Diagram of connections involved in dual-memory retrieval. A: Connections with single task practice only, before dual task. B: Model assuming a third intermediate task goal node. C: Model assuming associative chaining of response nodes. K = keypress task goal; V = vocal task goal; L = left key response; 4 = vocal response of “4.”

the premotor system. At first, the impact of this association may simply be to prime the second response so that, although participants are still performing the task through the two original, independent pathways, RT2 is reduced to some extent, perhaps falling below the sequential model prediction. Eventually, this new response association may strengthen to the point at which retrieval of the first response may lead directly to retrieval of the second response, with no mediation through the original pathway for the second task required. Thus, subsequential retrieval RTs may eventually be observed, even though retrieval is again mediated through only one set-cue node on a given trial. Note, however, that this account is viable only for items on which the vocal response is retrieved first (though not necessarily executed first). In the task used in this study, the vocal response uniquely determined the required keypress response, but the keypress response did not uniquely determine the vocal response.

The unification of the two responses into one retrievable unit may also be understood in terms of categorization processes. Rohrer et al. (1998) investigated when memories can and cannot be retrieved together. Their results suggested that within-category retrievals may be parallel but that across-category retrievals are generally sequential, perhaps because retrieval from alternating between categories requires a task goal change for every retrieval. Presumably, the participant must restart the retrieval process each time the task goal changes (for related work, see Rogers & Monsell, 1995). It may be possible to map our experiments onto those of Rohrer et al. It is plausible to assume that vocal and keypress task responses were represented as separate categories throughout single-task practice in our Experiment 1. As an example, one member of the category "vocal-digit" might be the stimulus response pair "blue → 4." One member of the category "keypress" might be "blue → left." With dual-task practice, however, a new category may be formed, for example, "blue → 4, left." In this case, "4" and "left" might be considered as exemplars of the category label, "blue." Perhaps this new category is formed only when participants group their responses. This category would be represented as a lone task goal to perform both tasks on the same trial. Once the dual-task responses have been categorized into a single-task goal, parallel retrieval of the complex dual response may be possible. This model may be compatible with the "new node" account mentioned earlier (see Figure 14B). Critically, the two responses in the dual task must be represented as a single, unified response if they are to be activated simultaneously.

Our results appear to pose some difficulties for certain models of memory and skill acquisition, such as the instance theory of automaticity (Logan, 1988, 1992), the extension of that theory by Nosofsky and Palmeri (1997; Palmeri, 1997), and the parallel retrieval model preferred by Wenger (1999). These models assume that a multi-step algorithmic process can be executed in parallel with, and independently of, direct and automatic retrieval of the answer. However, many algorithms discussed in the literature (e.g., arithmetic algorithms) are very likely to themselves consist, at least in part, of one or more memory retrievals (e.g., Logan, 1988; Compton & Logan, 1991; Palmeri, 1999; Rickard, 1997; Siegler, 1988). Thus, the instance theory assumption of parallel strategy execution appears to require that two memory retrievals from a single cue be completed in parallel. Our results appear to falsify such a prediction, even under the conditions in which single-task practice is substantial. This condition may mirror that of algorithms used in arithmetic and many other tasks, because the steps of such algorithms are likely to be highly practiced. Consider, for example, an algorithm of adding by 7s to calculate the answer to 4×7. Our results imply that even a simple calculation such as $7 + 7$ cannot take place in parallel with direct retrieval of the answer, 28, from memory (see Rickard, 1997, 1999, for empirical evidence favoring this hypothesis in the case of tasks such as arithmetic).

Earlier, we introduced the possibility that concurrent presence of both responses at the focus of working memory is necessary for practice to yield dual-task performance similar to that expected by parallel retrieval models. So far, we have treated this possibility solely as a framework within which our results might be understood. However, it also has potentially important implications in its own right for advancing the understanding of the role of working memory in associative learning. The IRI differences for groupers (less than 200 ms) and nongroupers (more than 400 ms) were only about 300 to 400 ms, on average, in Experiment 2. Thus, even for nongroupers, one would expect at least moderate levels of activation in working memory for the first response as the second response is executed. Yet, apparently for nongroupers it was at best very difficult to form any functional new associations between tasks, despite substantial practice. This finding raises the intriguing possibility that interitem (or interresponse) associative learning, whatever specific form it may take, cannot occur unless the two items are directly attended to within about 200–300 ms of each other. If correct, such a principle would place an important new constraint on the role of working memory in associative learning.

Another candidate account, which may be true either in addition to or instead of the temporal constraint account outlined, is that associative learning in tasks such as these depends crucially on the task goal (or task set) of the participant. It is possible that, for nongroupers, the goal for the first retrieval is inhibited after the first response has been retrieved. This goal inhibition may somehow shield the first task from subsequent learning on that trial, even though its response may still be active to some degree in working memory. The result would essentially be independent learning for each retrieval task on dual retrieval trials. On the other hand, groupers presumably have a task goal of executing both responses at the same time. As such, the goal-driven associative shielding just hypothesized would not occur, and some form of interitem associative learning, or chunking, might be expected. If this hypothesis proves correct, then the presence of an item in working memory would satisfy a presumably necessary, but certainly not sufficient, condition for associative learning involving that item. Future research exploring these possibilities is warranted.

References

Ashby, F. G., & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review, 93,* 154–179.

Carrier, L. M., & Pashler, H. E. (1995). Attentional limits in memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1339–1348.

Colonius, H. (1990). Possibly dependent probability summation of reaction time. *Journal of Mathematical Psychology, 34,* 253–275.

Compton, B. J., & Logan, G. D. (1991). The transition from algorithm to retrieval in memory-based theories of automaticity. *Memory & Cognition, 19,* 151–158.

Craik, F. I. M., Govoni, R., Naveh-Benjamin, M., & Anderson, N. D. (1996). The effects of divided attention on encoding and retrieval processes in human memory. *Journal of Experimental Psychology: General, 125,* 159–180.

Crutcher, R. J., & Ericsson, K. A. (2000). The role of mediators in memory retrieval as a function of practice: Controlled mediation to direct access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1297–1317.

Diederich, A., & Colonius, H. (1987). Intersensory facilitation in the motor component? A reaction time analysis. *Psychological Research, 49,* 23–29.

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

Gottsdanker, R. (1975). The attaining and maintaining of preparation. In P. M. Rabbitt & S. Dornic (Eds.), *Attention and performance V* (pp. 33–49). New York: Academic Press.

Li, Y., & Wright, D. L. (2000). An assessment of the attention demands during random and blocked practice schedules. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 53*(A), 591–606.

Logan, G. D. (1978). Attention in character classification tasks: Evidence for the automaticity of component stages. *Journal of Experimental Psychology: General, 107,* 32–63.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95,* 492–527.

Logan, G. D. (1992). Attention and preattention in theories of automaticity. *American Journal of Psychology, 105,* 317–339.

Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology, 14,* 247–279.

Nino, R. S. (2001). *Grouping two memory retrieval outputs from a single cue.* Unpublished manuscript, University of California, San Diego.

Nosofsky, R. M., & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review, 104,* 266–300.

Palmeri, T. J. (1997). Exemplar similarity and the development of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 324–354.

Palmeri, T. J. (1999). Theories of automaticity and the power law of practice. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 543–551.

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

Pashler, H. (1998). *The psychology of attention.* Cambridge, MA: MIT Press.

Pashler, H., & Christian, C. (1996). *Bottlenecks in planning and producing manual, vocal, and foot responses.* Unpublished manuscript, University of California, San Diego.

Pashler, H., & Johnston, J. C. (1989). Chronometric evidence for central postponement in temporally overlapping tasks. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 41*(A), 19–45.

Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin, 86,* 446–461.

Richardson, J. T. E. (1998). The availability and effectiveness of reported mediators in associative learning: A historical review and an experimental investigation. *Psychonomic Bulletin and Review, 5,* 597–614.

Rickard, T. C. (1997). Bending the power law: A CMPL theory of strategy shifts and the automatization of cognitive skills. *Journal of Experimental Psychology: General, 126,* 288–311.

Rickard, T. C. (1999). A CMPL alternative account of practice effects in numerosity judgment tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 532–542.

Rickard, T. C., & Bajic, D. (in press). Memory retrieval given two independent cues: Cue selection or parallel access? *Cognitive Psychology.*

Rickard, T. C., & Logan, G. D. (2002). *Memory retrieval practice: Strengthening or instance accrual?* Manuscript in preparation.

Rickard, T. C., & Pashler, H. (2003). *A bottleneck in memory retrieval from a single cue.* Manuscript submitted for publication.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124,* 207–231.

Rohrer, D., Pashler, H., & Etchegaray, J. (1998). When two memories can and cannot be retrieved concurrently. *Memory & Cognition, 26,* 731–739.

Ross, H. B., & Anderson, J. R. (1981). A test of parallel versus serial processing applied to memory retrieval. *Journal of Mathematical Psychology, 24,* 182–233.

Scarborough, D., & Sternberg, S. (1998). Discovering mental processing stages: The method of additive factors. In D. Scarborough & S. Sternberg (Eds.), *Methods, models and conceptual issues: An invitation to cognitive science* (Vol. 4, pp. 703–863). Cambridge, MA: MIT Press.

Schubert, T. (1999). Processing differences between simple and choice reactions affect bottleneck localization in overlapping tasks. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 408–425.

Schweickert, R. (1983). Latency network theory: Scheduling of processes in sentence verification and the Stroop effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 353–383.

Siegler, R. S. (1988). Strategy choice procedures and the development of multiplication skill. *Journal of Experimental Psychology: General, 117,* 258–275.

Sipos, I. (1982). RTs to colors and their names with a manual and a voice key. *Studia Psychologia, 24,* 107–144.

Townsend, J. T. (1990). Truth and consequences of ordinal differences in statistical distributions: Toward a theory of hierarchical inference. *Psychological Bulletin, 108,* 551–567.

Van Selst, M., Ruthruff, E., & Johnston, J. C. (1999). Can practice eliminate the psychological refractory period effect? *Journal of Experimental Psychology: Human Perception and Performance, 25,* 1268–1283.

Van Zandt, T. (2000). How to fit a response time distribution. *Psychonomic Bulletin & Review, 7,* 424–465.

Van Zandt, T., & Townsend, J. T. (1993). Self-terminating versus exhaustive processes in rapid visual and memory search: An evaluative review. *Perception & Psychophysics, 53,* 563–580.

Wenger, M. J. (1999). On the whats and hows of retrieval in the acquisition of a simple skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 1137–1160.

(*Appendix follows*)

Appendix

Perceptual Correction for the Sequential Model

The correct equation for RT2 according to the parallel model (Equation 2) is $\mu_{\mathrm{RT2}} = \mu_p + \mu_{\mathrm{prep}} + \mu_{[\max(\mathrm{RT}kr+\mathrm{RT}km,\mathrm{RT}vr+\mathrm{RT}vm)]}$. However, when single-task data are used to estimate this prediction, the resulting population equation (Equation 4) is biased: $\mu_{\mathrm{RT2}} = \mu_{\mathrm{prep}} + \mu_{[\max(\mathrm{RT}p+\mathrm{RT}km+\mathrm{RT}kr,\mathrm{RT}p+\mathrm{RT}vm+\mathrm{RT}vr)]}$.

Although a close approximation to the correct prediction, Equation 4 treats the single perceptual event of the dual tasks as two independent perceptual events, one for each task operating in parallel, rather than as a single event common to both retrievals, as assumed by the model. We estimated the magnitude of this bias through the following simulation. Because μ_{prep} is common to both of these equations, this component was ignored. Empirically motivated assumptions about the remaining components of RT2 were made. The value μ_p was assumed to be 100 ms, the combined retrieval and motor components for the keypress task ($\mu_{kr} + \mu_{km}$) were assumed to be 500 ms, and the combined retrieval and motor components for the vocal task ($\mu_{vr} + \mu_{vm}$) were assumed to be 700 ms. These values were reasonable given the single-task RTs observed in both experiments.

Distributions were assumed to be convolutions of a normal and an exponential distribution (i.e., the "ex-Gauss" distribution, with parameters μ, σ, and τ, where τ is the parameter for the exponential component) were assumed for all RT components. The ex-Gauss distribution has been shown to provide a good fit to RT data across a variety of tasks, although we make no claim that it is exact here. Estimates for the standard deviations for each of the RT components just mentioned were calculated under the assumption that the coefficient of variation (CV), σ/μ, equals 0.1 in the first simulation and 0.2 in the second simulation. Therefore, in the first simulation σ had a value of $0.1(\mu)$, and in the second simulation it had a value of $0.2(\mu)$. The parameter τ was also assumed to have values of $0.1(\mu)$ in the first simulation and $0.2(\mu)$ in the second.

Both the σ and τ estimates were based on item-level ex-Gauss fits to data from a simple letter naming task (Rickard & Logan, 2002). Participants received 100 trials for each of 10 letters presented in the middle of the computer screen with the same font, apparatus, subject pool, and a voice-key response as in the current experiments. We assumed that the total variability and skew of the perceptual, motor, and response mapping components of that task, relative to its mean RT, were similar to those of the current tasks. Beyond the first trial for each item, there was no RT speedup with practice. For data beyond the first block, item-level means (averaging only over trials for each item and each participant) averaged about 350 ms, and the estimated CV was slightly below 0.1. This approach to estimating CV is not confounded by variance between items, which was significant. To be conservative, we used CV values of both 0.1 and 0.2 in the simulations reported here. The estimate of the parameter τ also had an average value of less than $0.1(\mu)$ for the naming task. Again, to be conservative, we set values for τ to $0.1(\mu)$ and $0.2(\mu)$ in Simulations 1 and 2, respectively.

In each simulation, 1,000 simulated data points were generated for each of the equations described earlier. For example, a single simulated data point corresponding to Equation 2 was generated by obtaining a single random observation from the perceptual latency distribution and adding it to the maximum of one observation drawn from the keypress retrieval plus motor distributions, ($\mu_{kr} + \mu_{km}$), and one observation drawn from each of the vocal retrieval plus motor distributions, ($\mu_{vr} + \mu_{vm}$). A single simulated data point corresponding to Equation 4 was generated by obtaining two random observations from the perceptual latency distribution, adding one of these observations to a random observation from the keypress retrieval plus motor distribution ($\mu_{kr} + \mu_{km}$), adding the other observation to a random observation drawn from the vocal retrieval plus motor distribution, ($\mu_{vr} + \mu_{vm}$), and then selecting the greater of the two sums.

With σ and τ set to $0.1(\mu)$, the difference between the mean of the 1,000 simulated observations from Equation 2 and the mean of the 1,000 observations from Equation 4 was 0.72 ms. With σ and τ set to $0.2(\mu)$, the difference between these means was 12.8 ms.

Received July 12, 2001
Revision received October 22, 2002
Accepted October 24, 2002 ■