Psychological Research (1999) 62: 154–164

ORIGINAL ARTICLE

Pierre Jolicœur · Roberto Dell'Acqua

Attentional and structural constraints on visual encoding

Received: 20 July 1997 / Accepted: 17 March 1998

Abstract We examined the mechanisms that mediate the transfer of information from visual input to storage in memory. Observers performed two concurrent tasks, one of which required input into memory. We discovered that the processes involved in the transfer of information from sensory input into memory cause slowing in concurrent cognitive tasks (dual-task slowing). We used the dual-task slowing effect to demonstrate that memory encoding requires more time when more information is to be encoded and to show that dual-task slowing occurs long after the initial perceptual encoding of visual information (Exp. 1). These results suggest a late and central locus of interaction between the two tasks. Experiment 2 also used two concurrent tasks. However, we reversed the direction of interaction between them and produced a memory deficit from the execution of a concurrent task. Together the results suggest that the mechanisms that encode information into memory belong to a family of mechanisms that are involved in dual-task slowing phenomena and that have been studied under the rubric of the PRP effect (psychological refractory period). We were able to locate the most probable locus of the dual-task interactions to a process that appears necessary for memory encoding. We call this process *short-term consolidation*.

Introduction

In this article we focus on a key aspect of visual cognition, namely on the interface between vision and cognition. Visual cognition is usually studied in the laboratory by asking observers to perform specific tasks on well-controlled visual stimuli. Decision processes and other cognitive operations must be taken into account in order to understand and interpret the observed behaviour even for tasks that appear to depend primarily on visual perception, (e.g., Green & Swets, 1974).

Our present work examines the memory requirements of tasks in which the observer is asked to perceive simple visual patterns that are presented very briefly. Although performance in such tasks may appear to reflect primarily perceptual or sensory processing, especially when the amount of information to be reported is small (e.g., one or two letters), we argue that significant constraints on performance in many such tasks can be found at the point where perceptually-encoded information must be transferred to a short-term memory buffer. There is good evidence that perceptual encoding can take place rapidly and with relatively low capacity demands on other concurrent cognitive activities, but that the representations activated by early perceptual analysis are very short-lived (Potter, 1976, 1993). Although the information extracted from the visual display often must guide a response to be performed only one or two seconds after the termination of the visual display, this delay is long enough for the information to be lost unless it is transferred to a more durable form of memory. The short duration of early perceptual representations necessitates a transfer of information to a more durable form of memory, which we will call *durable storage*. We studied the capacity demands of this transfer process, which we call *short-term consolidation*.

P. Jolicœur (✉)
Department of Psychology,
University of Waterloo, Ontario, Canada N2L 3G1;
Tel.: 519-885-1211; Fax: 519-746-8631;
e-mail: pjolicoe@cgl.uwaterloo.ca;
Web: http://www.cgl.uwaterloo.ca/˜pjolicoe

R. Dell'Acqua
CREPCO, Aixen Provence

Experiment 1

The goal of this experiment was to demonstrate dual-task interference effects associated with memory encoding in a simple visual-encoding task. Subjects performed two tasks, one of which required visual-encoding and a deferred response. Therefore, this task also required storage of the information in memory. The other task provided a probe that we used to monitor the involvement of central processes in the visual-encoding task. For ease of exposition, we will proceed directly to describing the method and procedure used in Exp. 1, and then the results. In the subsequent sections, we will discuss the theoretical implications of these results.

Method

Subjects. The subjects were ten undergraduate students (five females) who volunteered to participate for pay or for course credit. All reported having normal or corrected-to-normal vision and normal hearing.

Visual stimuli. This stimuli were black characters presented on a white background on a color computer screen (CRT) controlled by a 486 CPU. Each character subtended .85° (height) × .8° of visual angle. When more than one character was displayed, the space between adjacent characters was .1°. The mask consisted of superimposed 0 and $ characters. The characters could be either upper-case letters or digits, randomly selected, without replacement, on each trial. The letters were selected from the set of consonants excluding S and Z, and the digits were selected among the digits 1 to 9, excluding 0.

Auditory stimuli. The auditory stimuli were pure tones with a frequency of 400 Hz or 1200 Hz. The stimuli were presented by the speaker on the monitor, and were well above threshold.

Procedure. A schematic representation of the trial structure is illustrated in Fig. 1, Panel a. The visual display contained either one or three characters. Each trial began with the presentation of a fixation box at the center of the screen. The box subtended 2.3° × 1.2° of visual angle for one-character displays, or 2.3° × 3.6° for three-character displays. The fixation box was intended to indicate the location, approximate area, and number of characters of the upcoming visual stimulus. Subjects initiated each trial by pressing the spacebar of a computer keyboard. The fixation box disappeared upon trial initiation, and after 400 ms a 250-ms visual display containing either one or three characters was shown, followed by a 100-ms mask. The characters were either letters or digits. The task associated with the visual stimulus was to remember as many characters as possible, when they were letters (remember condition). When they were digits, the characters could be ignored (ignore condition)[1].

After a delay of 350, 500, 650, 800, 1200, or 1600 ms from the onset of the visual display (stimulus onset asynchrony, or SOA), a tone was presented for 100 ms. Subjects were instructed to respond immediately to the tone by pressing the A (high pitch) or the Z (low pitch) keys on a QWERTY keyboard with the middle and the index fingers of their left hand, respectively. Response times to the tone were measured from the onset of the tone to the keypress. The tone task was defined as the primary task, and both speed and accuracy were strongly emphasized. After the response to the tone, subjects had to either report the letters shown within the visual display by typing them on the keyboard or to press the spacebar when digits were displayed. Recall was unspeeded and only accuracy was stressed by the instructions. Each subject performed 48 practice trials followed by 12 blocks of 48 experimental trials (576 experimental trials). The type of characters displayed, the SOA, as well as the pitch of the tone were fully crossed, with an equal number of replications of each possible combination occurring in a random order within each block of trials.

Results and discussion

Each trial produced one response to the tone and one response to the visual display. Only trials in which the response to the tone was correct were included in the analyses. The trials were then screened for outliers using a modified version of the procedure described by Van Selst and Jolicœur (1994)[2]. This procedure resulted in a loss of 3.2% of the correct trials. When an error was found in the auditory task, the entire trial was discarded. The results from both tasks were analysed using the analysis of variance (ANOVA) in which number of characters displayed, SOA, and task (remember vs. ignore) were treated as within-subject factors.

The most interesting results involve the response times in the auditory task for each SOA, depending on the number of characters displayed, and on whether the characters had to be remembered (letters) or ignored (digits). These results are shown in Panel a of Fig. 2. The results were clearcut and striking. First, response times to the tone were strongly elevated in the remember condition compared with those in the ignore condition, $F(1, 9) = 26.5$, $MSE = 17695$, $p < .001$. Second, the slowing effect was more pronounced in the three-character condition than in the one-character condition, $F(1, 9) = 28.2$, $MSE = 4189$, $p < .001$.

Third, there was more dual-task slowing at the shortest SOAs than at longer SOAs, $F(5, 45) = 21.6$, $MSE = 2424$, $p < .001$. There was also a highly significant three-way interaction between SOA, number of characters, and task, which is clearly evident in Fig. 2, Panel a, $F(5, 45) = 3.7$, $MSE = 1556$, $p < .007$. In the

[1] In several subsequent experiments, we counterbalanced the materials used in the remember and ignore conditions, such that the nature of the material was not confounded with whether the information had to be remembered or not. The results of these additional experiments were just as those for the present experiment, and they show that the differences between the remember and ignore condition were not due to the fact that one condition used letters and the other digits (Jolicœur & Dell'Acqua, 1998).

[2] In the Van Selst and Jolicœur (1994) procedure, the data in each cell are sorted, and the very largest observation is temporarily excluded from consideration. In the modified procedure, the observation that is furthest from the mean is temporarily excluded instead. (For distributions with positive skew, the outcomes will be very similar.) The mean and standard deviation of the remaining numbers is then computed. Cutoff values are established using the following equations:

$$V_{low} = \overline{X} - C * SD \quad (1)$$

$$V_{high} = \overline{X} + C * SD \quad (2)$$

The smallest and largest observation in the cell are then checked against the cutoff values, V_{low} and V_{high}. If one or both are outside the bounds, then they are defined as outliers and excluded from further consideration. If an outlier is found, then the algorithm is applied anew to the remaining data. The value of C depends on the sample size such that the estimated final mean is not influenced by sample size. See Van Selst and Jolicœur (1994) for additional details.

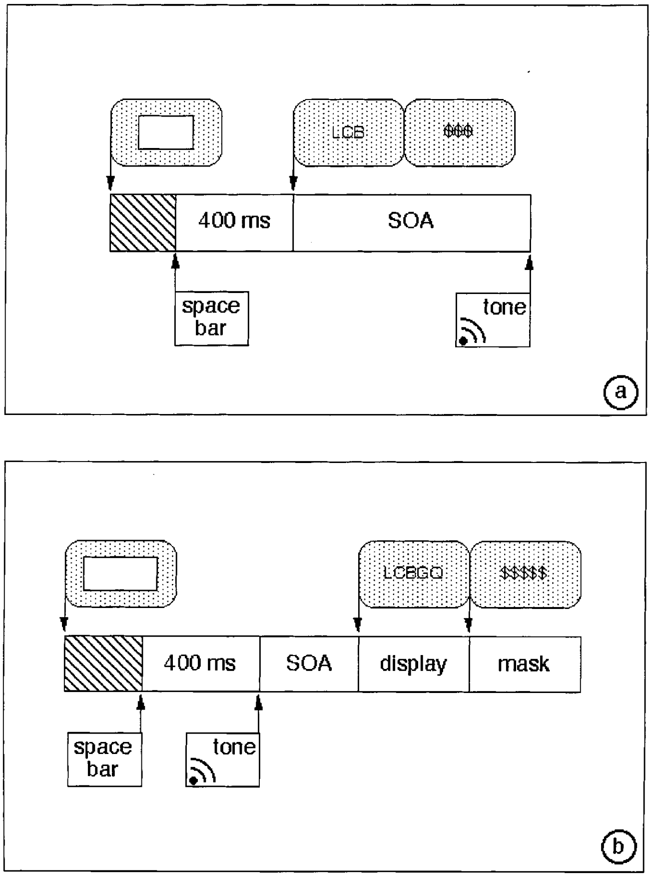

Fig. 1 Experimental paradigms. **Panel a** Exp. 1. For explanation, see text. **Panel b** Exp. 2. For explanation, see text

remember condition (letters), there was a larger effect of SOA for the three-letter condition than for the one-letter condition. In marked contrast, in the ignore condition (digits), the effects of SOA across the one-digit and three-digit conditions were identical. A separate analysis performed on the data from the ignore condition revealed a significant effect of the SOA, which was confined only to the first portion of the SOA function (SOA = 350 ms), $F(5, 45) = 4.9$, $MSE = 1265$, $p < .002$. The separate analysis did not reveal any other significant effect. We discuss this interaction in the final part of the present section.

An analysis was also carried out on the mean proportion of accuracy in performing the auditory task as a function of the same variables as those considered in the analysis of the reaction times. The results are shown Fig. 2, Panel b. There was a slight reduction (.01) in the task accuracy in the ignore condition compared to the remember condition, $F(1, 9) = 4.1$, $MSE = .0005$, $p < .07$. A significant three-way interaction between SOA, task, and number of characters was also found and apparently suggested that this reduction in accuracy was due to the small drop in accuracy in the one-digit condition at the shortest SOA, which is evident in Fig. 2, Panel b, $F(5, 45) = 2.9$, $MSE = .001$, $p < .03$. Overall, the accuracy results allowed us to interpret the pattern of response times with no concern for speed-accuracy trade-offs.

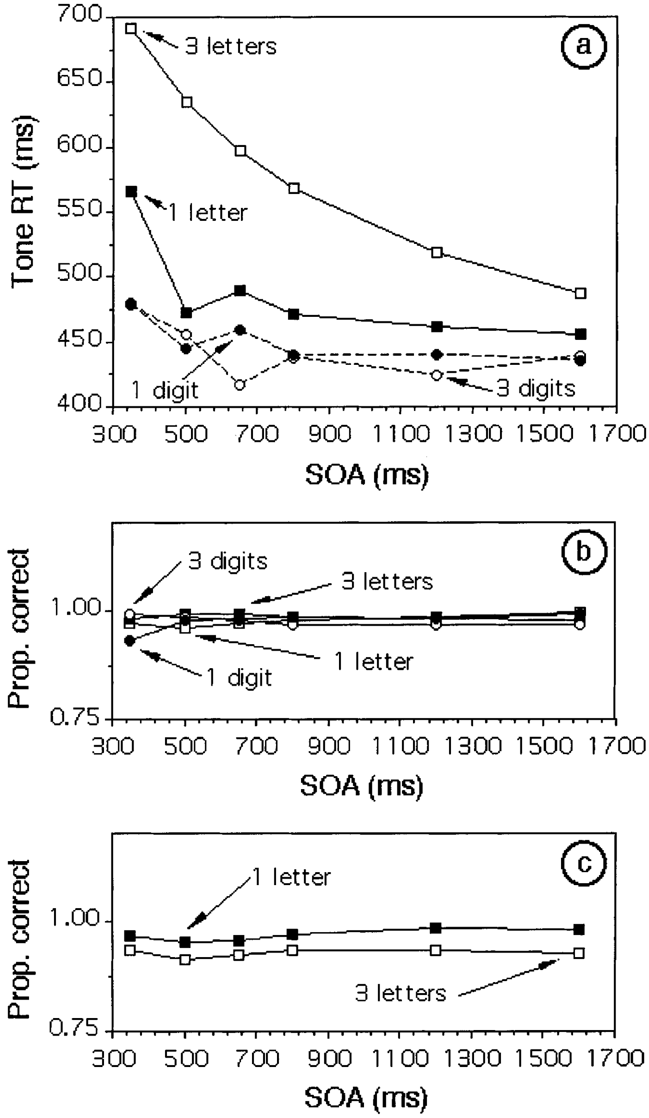

Fig. 2 Results of Exp. 1. **Panel a** Mean response times to the tone depending on characteristics of the visual display (1 vs. 3 symbols; letters vs. digits) and on the SOA between the visual display and the tone. **Panel b** Proportion of correct responses for the tone task for the same variables as in Panel a. **Panel c** Proportion of correctly reported letters, without regard to report order, for trials with 1 or 3 letters at each SOA

The mean proportion of letters correctly recalled at the end of the trial is shown in Fig. 2, Panel c. An analysis carried out on these results indicated a worse recall in the three-letter condition than in the one-letter condition, $F(1, 9) = 5.5$, $MSE = .008$, $p < .05$. Neither the main effect of SOA nor the interaction between SOA and number of letters were significant in this analysis ($F < 1.5$, $p > .2$ in both cases). As can be seen in Panel c, recall performance was essentially flat across SOAs, such that the amount of information recalled was not affected by the delay between the visual display and the onset of the tone. This result is in sharp contrast with the effects of SOA on response times to the tone (see Panel a).

We interpret the results of Exp. 1 in the following way. After encoding the visual display, the representations were evaluated to decide whether the information was to be remembered or discarded. In the ignore condition, the elevated response times at the shorter SOAs likely reflect the cost of deciding whether to remember or to discard the information. When discarded, no further processing of the visual information was required, so it did not matter whether one or three digits had been shown, and response times to the tone decreased to a common asymptote. In contrast, in the remember condition, additional processing was initiated. This processing caused an increase in response times to the tone, especially at shorter SOAs, even when only one letter had to be remembered. The increase in response times to the tone was larger when three letters had to be remembered. Furthermore, in this case the difference between the remember and ignore condition extended for a longer period of time. It is of the utmost theoretical importance to realize that the shortest SOA (350 ms) was long relative to the time required to encode 1–3 alphanumeric characters (Shibuya & Bundesen, 1988). Despite this long delay, response times to the tone were elevated in the remember condition relative to the ignore condition, even when only one item had to be remembered. This critical observation rules out an early locus of interference between the two tasks. We conclude that there is a process engaged by the memory task that occurs after the characters are encoded and that this process is associated with the slowing effect observed in the auditory task. This process is engaged when information is to be encoded (the remember condition), but not when it can be discarded (the ignore condition). Furthermore, this process requires more time to run to completion when more information is to be remembered (see Fig. 2, Panel a, three-letter vs. one-letter conditions).

We call the memory process revealed by the results of Exp. 1 *short-term consolidation* (STC). We hypothesize that information not subjected to the process of short-term consolidation will not be remembered. We tested this hypothesis by asking, on the last trial of several experiments like Exp. 1, for recall of information that was ignored at the beginning of the trial (e.g., asking for recall digits in the present case). Performance was at chance, suggesting that explicit recall requires short-term consolidation (Jolicœur & Dell'Acqua, 1998).

In addition to short-term consolidation, two additional processes producing dual-task slowing are suggested by the results. The first is associated with the decision to remember or discard the information in the visual display. The elevated response times in the ignore condition in the 350-ms SOA condition reflect this decision process. The additional slowing in the one-letter condition relative to that in the one-digit condition at 350-ms SOA reflects the additional cost associated with short-term consolidation for one item. In other experiments, we required the subject to remember the characters in the visual display on every trial. This eliminated the decision to remember or discard the visual information. Considerable dual-task slowing was still observed, with a similar effect of the amount of information to be remembered as in Exp. 1. Thus, remembering the information per se is sufficient to produce dual-task slowing. The second additional source of slowing appears to be the retention of information in memory. However, this effect was small, amounting to about 24 ms when one item had to be remembered (see Panel a of Fig. 2). The results suggest that the memory-load effect was larger when three items had to be remembered. However, it appears that response times to the tone had not reached asymptote even at the longest SOA, suggesting that the difference between RT for the three-letter condition and RT for digits at the longest SOA is probably an overestimate of the cost of retention. The results suggest strongly that the cost of retention per se was much smaller than the cost of short-term consolidation, even in the three-letter condition.

Recall was very high both for the one-letter and the three-letter conditions, and there was no effect of SOA on performance, as can be seen in Panel c of Fig. 2. In contrast to the response times in the auditory task, the retention of the information was not differentially affected by when we presented the tone.

A model of task interactions

Figure 3 (Panels a–c) present a model of hypothesized interactions between the processes required to perform the two tasks executed in Exp. 1. In the model, certain cognitive operations require processes that are mediated by brain mechanisms that have limited bandwidth. These mechanisms act as a single channel in the flow of information processing. If two concurrent tasks both require the single channel, an information processing bottleneck results, and processing for one of the two tasks must wait. The waiting period ends only when processing of the information for one task clears the bottleneck stage or stages such that the single channel is now available for the other task. See Pashler (1994) for a review of evidence supporting this type of model, and Meyer and Kieras (1997) for some diverging opinions.

Although certain operations must be performed in a serial fashion across tasks, other operations can be performed in parallel, as schematized in Panel a. We call the first stage of processing for the memory task *sensory encoding* (*SE* in Fig. 3), and we hypothesize that representations at this level are susceptible to masking. Sensory encoding provides input to in the next level, *perceptual encoding* (*PE* in Fig. 3). As in Duncan (1980), we hypothesize that perceptual encoding produces representations that include information about both sensory (e.g., colour, size) and more abstract properties (e.g., letter identity). Although we postulate the output

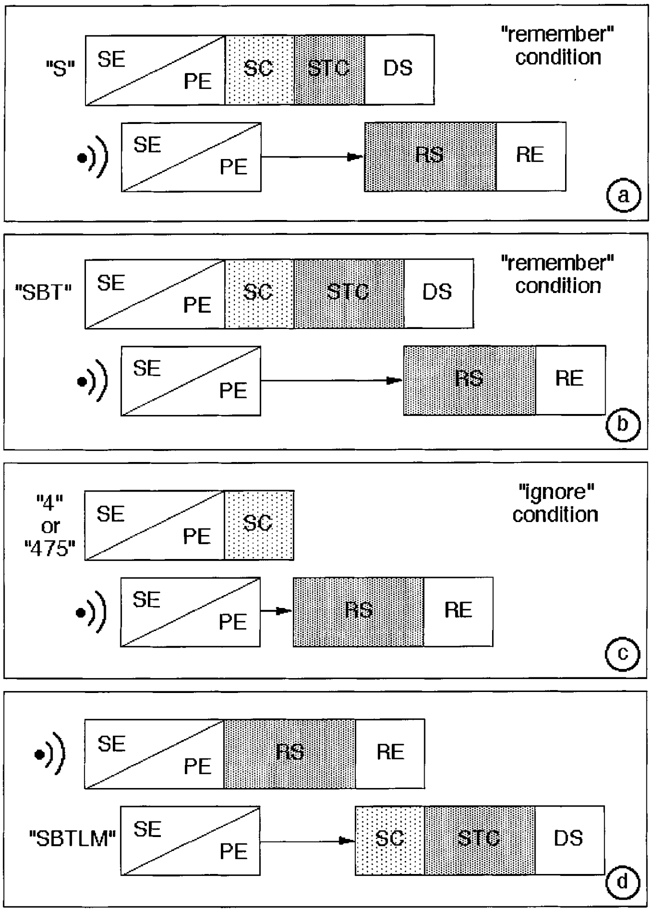

Fig. 3 Model of postulated interactions between stages of processing in the two tasks. The memory task requires sensory encoding (*SE*), perceptual encoding (*PE*), selective control (*SC*), short-term consolidation (*STC*), and durable storage (*DS*). The auditory task requires SE and PE, followed by response selection (*RS*) and response execution (*RE*). **Panels a–c** Exp. 1. **Panel a** Remember condition, 1 letter, short SOA: selective control (*SC*) and short-term consolidation (*STC*) postpone response selection (*RS*), elevating RT to the tone. **Panel b** A longer period of short-term consolidation (*STC*) results when 3 letters are memorized, which postpones response selection (*RS*) for the tone for longer than when 1 letter (Panel a) is memorized. **Panel c** In the ignore condition, the information is not subjected to short-term consolidation (*STC*); only selective control (*SC*) postpones response selection (*RS*). **Panel d** Exp. 2. Response selection (*RS*) in the auditory task postpones selective control (*SC*) and short-term consolidation (*STC*) in the memory task. Because the visual display is masked, the output of perceptual encoding (*PE*) decays if not subjected to short-term consolidation (*STC*). A longer period of postponement results when SOA is reduced, leading to lower recall

of perceptual encoding is no longer maskable by sensory stimuli, we hypothesize that these representations decay unless processed further (Potter, 1993).

In Exp. 1, a decision had to be made between remembering the information and ignoring it. We call the processes that performed this decision and controlled subsequent processing *selective control* (*SC* in Fig. 3). In the remember condition, selective control was followed by a process of short-term consolidation (*STC* in Fig. 3), which we believe is necessary for later conscious report of the information. The end result of short-term consolidation is a durable form of memory, durable storage (*DS* in Fig. 3), not subject to rapid decay.

In the auditory task, following early encoding, response selection (*RS* in Fig. 3) is engaged. A substantial body of evidence implicates response selection as one of the operations that require single-channel mechanisms. Experiment 1 suggests that both selective control and short-term consolidation in the memory task can postpone the response selection in the auditory task. In the remember condition, response selection for the auditory task must wait for short-term consolidation to finish. A shorter period of waiting results when one letter is to be remembered (Panel a) than when three letters are to be remembered (Panel b), because short-term consolidation takes less time for one letter than for three. When the visual information could be ignored, only selective control postponed response selection in the auditory task (Panel c). What is new about our present results and formulation is not that memory encoding is somehow capacity-limited (Sperling, 1960), but rather it is the demonstration that memory encoding requires the same family of bandwidth-limited mechanisms as those required to perform response selection.

Experiment 2

If response selection and short-term consolidation both require access to the same limited-capacity central system, then it should be possible to interfere with short-term consolidation by occupying the central system with operations required to perform a concurrent task. In a visual-encoding task, this interference would manifest itself as a reduced ability to report the information that was presented. We tested this hypothesis in Exp. 2. We modified the experimental paradigm to allow the auditory task to engage the limited-capacity channel first. The information for the memory task (visual input) was masked, so the output of perceptual encoding would decay unless subjected to short-term consolidation. The longer the postponement of short-term consolidation by response selection was, the greater the decay, and the worse the recall of the visual information. These relationships are diagramed in Panel d of Fig. 3. In Exp. 2, we used the same auditory tasks as in Exp. 1, but the tone was presented before the visual display, that consisted of five letters (masked). The visual task was to recall as much information as possible at the end of the trial.

The model in Panel d of Fig. 3 leads us to make two predictions. First, recall should decrease as the SOA is decreased. Representations created or activated by perceptual encoding that are not subjected to short-term consolidation decay rapidly in the absence of bottom-up support from sensory input. This bottom-up support is terminated with the onset of the pattern mask. Decreasing the SOA increases the likelihood that short-term consolidation will have to wait by increasing the likelihood that response selection in the auditory task still occupies the limited-capacity channel when the mask limits bottom-up support for representations produced by perceptual encoding.

The second prediction is that there should be a dependency between response time in the auditory task and recall at short SOA. On trials in which the auditory task was executed rapidly (shorter response times), the mean duration of response selection is more likely to have been shorter than on trials in which the auditory task was executed slowly (longer response times). The implication is that for shorter response times, the limited-capacity channel will have been released earlier. In this view, the delay of the onset of short-term consolidation would be smaller, which would reduce the amount of decay of representations activated by perceptual, which in turn should result in better recall.

Method

Subjects. The subjects were 24 undergraduate students (10 females) who volunteered to participate for pay or for course credit. All reported having normal or corrected-to-normal vision and normal hearing.

Visual stimuli. The visual stimuli were composed of strings of five upper-case black letters presented on a white background on a colour computer (CRT) screen controlled by a 486 CPU. The dimensions of the letters and the space between adjacent letters was the same as in Exp. 1. Each string subtended 5.15°. The letters were randomly selected without replacement from the set of consonants, excluding S and Z. The mask was composed of five $s shown at the same location as the letters.

Auditory stimuli. The auditory stimuli were pure tones with the same frequencies as those used in Exp. 1.

Procedure. A schematic representation of the trial structure is illustrated in Fig. 1, Panel b. A centrally displayed fixation box (6 × 2.3°) indicated the location and area of the upcoming visual stimulus. Subjects initiated each trial by pressing the spacebar of a computer keyboard. The fixation box then disappeared, and after 400 ms a 100-ms high-pitch or low-pitch tone was presented. As in Exp. 1, subjects were instructed to respond immediately to the tone by pressing the A key (high pitch) or the Z key (low pitch) with the middle and the index fingers of their left hand, respectively. This task was defined as the primary task, and both the speed and the accuracy of this response were emphasized. At an SOA of 50, 200, 400, or 800 ms following the tone, a string of five letters was displayed for either 117 ms or 200 ms, followed by the mask of $s, which was exposed for 300 ms. Subjects were told to memorize the characters within the string and to report as many characters as possible by typing them on the keyboard in the order in which they appeared on the computer screen. Recall was unspeeded, and only accuracy was stressed by the instructions. Subjects were informed that the string was always composed of different letters.

Each subject performed two blocks of 32 practice trials followed by ten blocks of 32 trials (320 experimental trials). Exposure duration of the letter string, SOA, as well as the pitch of the tone were fully crossed, with an equal number of replications of each possible combination occurring in a random order within each block of trials.

Results and discussions

The outlier-screening procedure used in Exp. 1 was also used here, resulting in the loss of 2.9% of the correct trials. As in Exp. 1, when an error was found in the auditory task, the entire trial was discarded. The results from both tasks were analysed with an ANOVA in which exposure duration of the visual display and SOA were treated as within-subject factors.

The results are shown in Panels a–e of Fig. 4. Recall (Panel a) decreased as the SOA between the tone and the letters decreased, $F(3, 69) = 52.7$, $MSE = .022$, $p < .001$. Recall was higher for a 200-ms exposure duration than for 117 ms, $F(1, 23) = 228.3$, $MSE = .041$, $p < .001$. In addition, the effects of SOA were larger when the exposure duration was shorter, $F(3, 69) = 4.8$, $MSE = .023$, $p < .005$. We discuss this interaction below. At first glance, the effects of SOA might appear small (a loss of .33 letters from the longest to the shortest SOA). No one would doubt, however, the potency and importance of manipulations of exposure duration in the whole-report task. An increase of 83 ms of exposure duration produced a recall gain of .44 letters. If we use the magnitude of the exposure duration effect as a measure for that of SOA, it is obvious that the effects of SOA were, in fact, relatively quite large.

Recall performance was also analysed as a function of the speed in carrying out the auditory task. Recall was computed for each subject, in each cell, for trials with response times (RT_1s) above or below the median RT_1 in that cell, following outlier screening of the data set. In Panel b of Fig. 4, recall performance is shown conditionalized on the length of RT_1. At the shorter SOAs, recall was higher for shorter RT_1s than for longer RT_1s, $F(3, 69) = 9.0$, $MSE = .037$, $p < .001$. This is what we would expect if a central bottleneck was occupied for a longer time when RT_1 was long than when RT_1 was short, and if this longer occupation interfered with short-term consolidation. This effect, however, is only expected at shorter SOAs, when contention for bottleneck mechanisms is more likely,. At the longer SOAs (especially at 800 ms), we would not expect this effect, because the response to the tone has already been made on most trials (see Panel e). The diminishing difference in memory performance across trials with short versus long RT_1 as SOA was increased was, therefore, as expected from the model shown in Panel d of Fig. 3. As exposure duration was increased, more information was recalled. We assume this effect has an early locus and that the change in recall as function of exposure duration reflects the rate of early encoding processes. It is instructive to consider how the effects of other variables interact with those of exposure duration, which is shown in Panels c–d of Fig. 4. Suppose that response selection

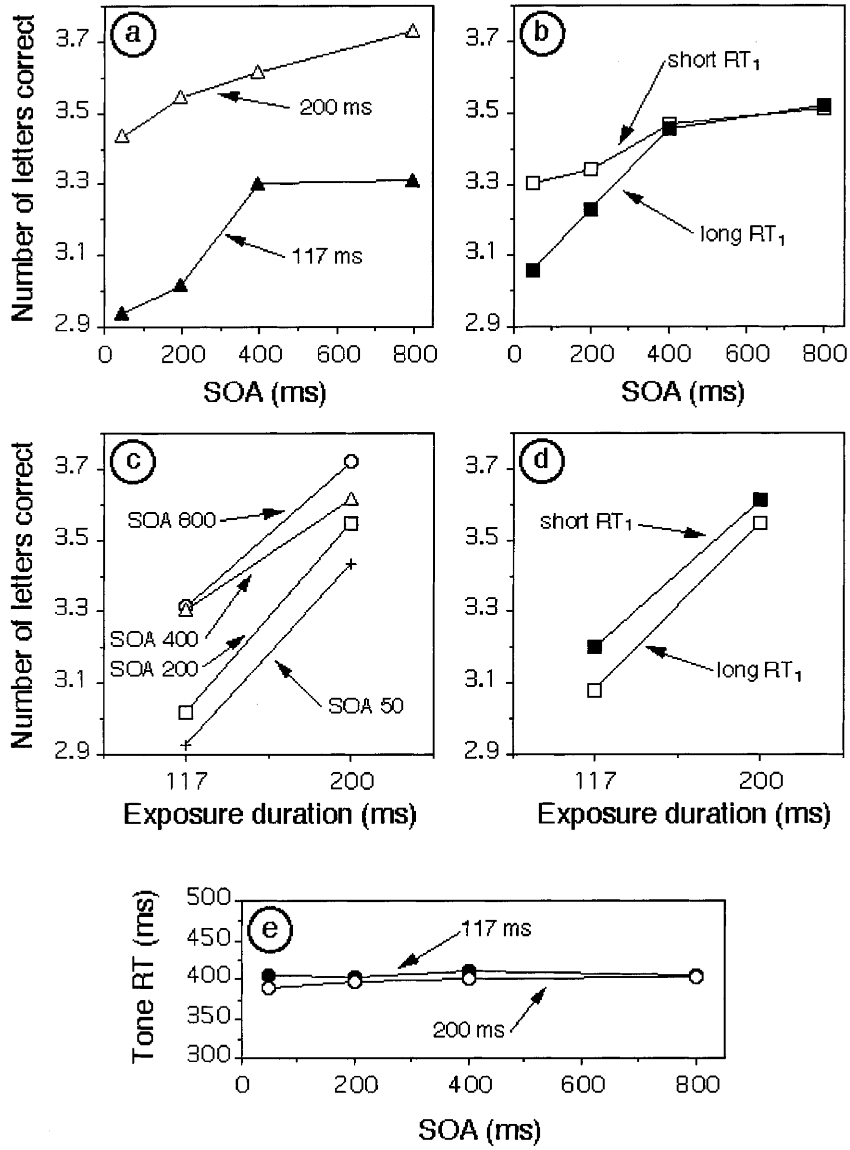

Fig. 4 Results of Exp. 2. **Panel a** Mean number of correctly-reported letters in the visual task (*max* = 5) without regard to order or position for each combination of exposure duration and SOA. **Panel b** Mean recall for each SOA conditionalized on response time for the tone (RT_1). **Panel c** Effects of exposure duration on recall for each SOA. **Panel d** Mean recall for each exposure duration conditionalized on response time to the tone (RT_1). **Panel e** Mean response times for the auditory task for each combination of exposure duration and SOA

interferes with perceptual encoding (rather than short-term consolidation, as we propose). An increase in interference would result in a lower rate of information extraction, whereas a decrease in interference would result in a higher rate. That is, the slope of the duration effect should be shallower when interference was higher, and steeper when interference was lower. Panel c shows recall as a function of exposure duration for each SOA. Although there was a significant interaction between duration and SOA, the pattern of results is not consistent with the view that the rate of information extraction slowed as SOA was shortened, because the steepest slopes for the duration effect are associated with the shorter SOAs, where the interference was greatest.

It would have been better for our theory if there had been no interaction between SOA and duration, on the view that duration has an early locus (sensory or perceptual encoding), whereas SOA is hypothesized to have a different and later locus (postponement of short-term consolidation). Further analyses in which we performed a median split of the subjects based on the their overall accuracy in the auditory task revealed the following pattern of results. For subjects with high accuracy, there was no hint of the interaction shown in Panels a and c. In other words, for subjects who performed the auditory task very accurately, the effects of SOA and duration were independent, as the theory would lead us to expect. In several subsequent experiments, we found this independence between SOA and duration, even when all the subjects were included in the analysis. We are not sure why or how the interaction occurred for the less accurate subjects in the present experiment. It is possible that on some trials they may have processed the letters at the expense of accuracy in the tone task. On the viable assumption that this occurred more frequently for the longer exposure duration, it could have produced the interaction between SOA and duration, which is not present in subjects with above-median accuracy.

Thus, our results are generally consistent with the view that early stages of processing have a large bandwidth and that capacity limitations arise later. Furthermore, Panel d of Fig. 4 shows the magnitude of the effect of exposure duration for different speeds of response in the auditory task. Recall was greater for short RT_1 than for long RT_1 (see also Panel b). Furthermore, the slope of the effect of duration on recall was not statistically different for short versus long RT_1 $F(1, 23) = 2.1$, $MSE = .040$, $p > .16$. Certainly, the magnitude of the exposure duration effect did not decrease as response times lengthened (since the slope was numerically larger).

Response times and accuracy in the auditory task were analysed as a function of the same variables considered for the analysis of recall performance. There was a small effect of exposure duration on RT_1s, with slightly longer RT_1s at the shorter exposure duration relative to the longer exposure duration of the visual display, but the magnitude of the effect was small. (The largest difference between any two means was less than 10 ms.) The mean accuracy in the auditory task (from the shortest SOA to the longest) was .93, .95, .96, and .96. The lower accuracy at the 50-ms SOA compared with that at the other three SOAs was significant, $F(3, 69) = 6.2$, $MSF = .0014$, $p < .001$. This suggested that we may have underestimated the interference on the memory task on the assumption that the loss of accuracy in the auditory task resulted from trials in which subjects processed the visual stimulus at the expense of the auditory one. There were no other significant effects, all ps $> .66$ in both cases.

General discussion

The results of both experiments demonstrate dual-task interference between a simple cognitive task – a speeded discrimination between two easily distinguished auditory signals – and tasks that depend primarily on encoding information into memory from visual input. In Exp. 1, when the information to be memorized was presented first, we observed interference form the memory task on the auditory task. Response times to the auditory signals were systematically slowed as the demands of the memory task were increased (Fig. 2, Panel a). This effect shows the characteristic increase in response times as SOA is decreased that is observed in the dual-task slowing literature. In Exp. 2, we presented the tone first and information to be memorized second. This allowed the processes required for the auditory task to gain access to central mechanisms first. As expected on this analysis, response times to the tone were not affected by the SOA between the tone and the visual display (Fig. 4, Panel e). However, now there was a decrease in the amount of information recalled as the SOA was reduced. We can model these dual-task interactions with the models illustrated in Fig. 3 (Panels a–d). A central hypothesis in these models is that the delayed report of information requires access to a durable form of memory, which we call durable storage (*DS*). Information can only be placed in durable storage if it undergoes a process of short-term consolidation (*STC*). However, short-term consolidation requires central mechanisms that are also required for response selection (*RS*). Thus, if the tone is presented first and response selection processes occupy central mechanisms, performance in the memory task suffers. Conversely, if short-term consolidation for the memory task is allowed to have access to central mechanisms first (the tone is presented second), then response selection for the tone must wait, which we observe as dual-task slowing.

Speed versus accuracy

Dual-task interference, which in our theory results from the postponement of a stage of processing, can manifest itself in the results in at least two ways. We demonstrated both manifestations in the two experiments in this article. The first was a slowing of response times in the auditory task in Exp. 1 in the absence of changes in error rates. Here there was a change in speed without a change in accuracy. The second was observed in the memory task in Exp. 2 as a decrease in recall (a change in accuracy) in a paradigm in which we did not measure speed (in the case of the memory task). In our view, these different manifestations of dual-task interference depend critically on the methods used to present the stimuli, and in particular on masking. In the absence of masking (e.g., the tones in the auditory task), persistence allows the perceptual system to bridge the period of time during which central mechanisms are busy with the other task. The result is a slowing of responses without a decrease in accuracy because persistence allowed the information in the stimulus to survive the period of postponement. In contrast, when the stimulus is masked, it is possible to observe dual-task interference as a loss of accuracy even if the task is unspeeded, as was the case in the memory tasks in Exp. 2.

The foregoing observations on the importance of masking can explain previous results which may appear at odds with the present findings. In particular, Blake and Fox (1969) asked subjects to report the identity of a briefly presented letter that was shown at various SOAs following the onset of a tone. Unlike us, however, they found no evidence for dual-task interference. There are several differences between their experiment and ours. For example, they adjusted the level of performance by manipulating exposure duration, in unmasked visual displays. Shorter durations were used to produce lower performance that would not be at ceiling. Interestingly, shorter stimulus durations often result in longer persistence, a result known as the inverse-duration effect (Coltheart, 1980). It is possible that persistence prevented the loss of accuracy in the Blake and Fox experiment. Evidence supporting the suggestion that the absence of a pattern mask in their paradigm was critical in producing the null effects of SOA was provided by Jolicœur (in press-c). He tested this hypothesis directly by performing experiments similar to those of Blake and Fox, with and without a pattern mask following the visual stimuli. Significant effects of SOA similar to those in Exp. 2 were found, but only when a pattern mask followed the visual targets. The effects of SOA were eliminated when the pattern mask was removed.

The present work is relevant for a complete understanding of a number of phenomena that involve encoding information into durable storage in the presence of concurrent processing demands. Consider a recent paper by Duncan, Ward, and Shapiro (1994). They presented two events staggered in time, each event consisting of an alphanumeric symbol followed by a pattern mask. When the second symbol followed the first by 100–300 ms, report of the identity of the second symbol was worse than at longer intervals. This task required encoding into durable storage (via short-term consolidation) because of the delayed responses. Thus, a sufficient account of their results is that the effects they observed were due to interference in the short-term consolidation of the second symbol. We hypothesize that short-term consolidation of one symbol can interfere with short-term consolidation of another (Chun & Potter, 1995). For this to occur, however, the symbols must be presented at different points in time, such that one gains access to short-term consolidation mechanisms first, before the second one is selected for short-term consolidation. Under these conditions, either short-term consolidation or the selection for short-term consolidation for the second symbol will be postponed. If the symbol are masked, then there will be a loss of information for the second symbol during the period of postponement due to the short-lived nature of representations produced by perceptual encoding that are not supported by ongoing bottom-up activation.

The results of the two experiments in this article and the theory we propose do not mesh well with previous results and associated theoretical interpretations produced by Pashler (1989, 1993). Pashler found small or nil effects of SOA in various visual-encoding tasks that followed immediately after an auditory discrimination task (as in Exp. 2 of this article). He interpreted these results as evidence for capacity-free visual encoding from early stages of encoding all the way to, and including entry into, durable storage. At the moment, we do not have a complete understanding of the experimental parameters and boundary conditions that lead to the diverging patterns of results. It is important to note that the effects that we report in this article are robust and easily replicated, and that we were able to extend them to related paradigms (e.g., Jolicœur, in press-a, b, c; Jolicœur & Dell'Acqua, 1998). Therefore, we have found conditions that reliably produce dual-task interference in visual-encoding tasks, and our results show that one or more stages of processing required to perform various visual encoding tasks are likely capacity-limited, at least under the conditions instantiated in our experiments.

What then could explain the small or null effects observed in Pashler's (1989, 1993) experiments? There are several possibilities. De Jong and Sweet (1994) have shown that relative task preparation can affect the outcome of experiments using Pashler's (1989) paradigm. De Jong and Sweet showed that increasing preparation for the tone task increased the size of SOA effects observed in the visual-encoding task. They argued and provided evidence that Pashler's (1989) subjects were relatively unprepared for the easy auditory task and more prepared for the difficult visual task, which would tend to minimize the effects of SOA on performance in the visual task. The second possibility is that the display characteristics and masking stimuli used by Pashler may have taken performance off the ceiling without eliminating the persistence of the information encoded from the display. Giesbrecht and Di Lollo (1998) have demonstrated that the type of masking used in dual-task visual-encoding experiments is critical. Masking by integration (Scheerer, 1973) degrades the identifiability of a visual target without altering persistence. Consistent with this model, Giesbrecht and Di Lollo found no effects of SOA when the second target in a dual-task encoding paradigm was masked by intergration (although performance was lowered significantly by the integration mask, showing that masking was indeed taking place). In contrast, large effects of SOA were found with masking by interruption (Scheerer, 1973). This possible account of the Pashler results does have difficulty with the fact that his display durations were relatively long and that masking by integration is less likely as the SOA between the display and the mask is increased (Scheerer, 1973). The main point here, however, is that performance that is off ceiling does not necessarily imply that persistence was eliminated. Coltheart (1980) argued that a distinction should be made between visible persistence and informational persistence. While visible persistence usually exhibits an inverse-duration effect, informational persistence, if anything, tends to exhibit the opposite pattern (i.e, longer persistence following a longer display). In our model (Fig. 3) it is possible that the long display duration used in Pashler (1989) could have produced a long period of information persistence in representations produced by perceptual encoding operations (*PE* in Fig. 3). Although these representations are not sufficient for overt responses (Duncan, 1980), their persistence could have bridged the period of time during which central mechanisms were busy with the tone task. Following the central processing required to perform the tone task, short-term consolidation could proceed if the results of perceptual encoding had not faded completely. Therefore, our theory can, in principle, explain both our positive results and Pashler's (1989, 1993) negative results on the assumption that there were different durations of information persistence across the paradigms in representations activated by perceptual encoding operations. This account should be amenable to empirical testing.

The theory outlined in Fig. 3 can also provide the basis for an account for other dual-task interference effects, such as the attentional blink phenomenon (e.g., Raymond, Shapiro, & Arnell, 1992; Broadbent & Broadbent, 1987). We suppose that the short-term consolidation of a first target can postpone the short-term consolidation of a second target for some time in a manner similar to what is diagramed in Fig. 3, Panel d (but with short-term consolidation in the first task taking the place of response selection). Several recent results from our laboratory provide converging support for this theory. Manipulations of the task to be performed on the first target ($Task_1$) that should affect the duration of operations believed to require central bottleneck mechanisms have been found to modulate the magnitude of the attentional blink effect in the expected way (Jolicœur, in press-d). The theory also predicts that the attentional blink effect should be observable using cross-modal stimuli. Although one experiment has failed to find significant cross-modal interference (Duncan, Martens, & Ward, 1997), three other laboratories have each reported several experiments in which a crossmodal attentional blink has been observed (Arnell & Jolicœur, in press; Jolicœur, in press-b, Potter, Chun, Banks, & Muckenhoupt, 1998; Shullman & Hsieh, 1995). Therefore, there is very good evidence that it is possible to observe a crossmodal attentional blink effect if appropriate conditions are met. As expected if persistence of information about the second target can bridge the period of postponement created by the processes required to carry out $Task_1$, there is no attentional blink if the second target is not masked (Giesbrecht & Di Lollo, 1998; see also Jolicœur, in press-c). Finally, if a speeded response is required in $Task_1$, then it can be shown that the magnitude of the attentional blink effect is smaller if RT_1 (response time in $Task_1$) is short than if RT_1 is long (e.g., Jolicœur, in press-a,b,d). This result is similar to the results shown in Fig. 4, Panels b and d. We believe that it has the same cause: the rapid decay of the representation produced by perceptual encoding (*PE*) of a masked second target during a period of central postponement of the short-term consolidation process required to encode the second target into durable storage (*DS*). This central postponement is longer when RT_1 is long than when RT_1 is short.

In conclusion, we focused on an important component of the processing and the cognitive architecture required to mediate performance in many paradigms in the area of visual cognition, namely on the interface between perception and short-term memory. Our results suggest that there are both attentional and structural constraint operating at the interface between perception and memory. The primary structural constraint is that encoding information into memory requires capacity-limited central brain mechanisms. The process required for input into durable storage is called short-term consolidation in our theory. Short-term consolidation is capacity limited and requires more time when more information is to be encoded. The primary attentional constraint is that short-term consolidation is subject to dual-task interference. We have proposed a theoretical account of this interference in which short-term consolidation is postponed by concurrent processing (such as response selection) that also requires the same capacity-limited central mechanisms. Jolicœur and Dell'Acqua (1998) describe several additional experiments investigating the properties of short-term consolidation, and they present computer simulations and additional arguments in support of the theory.

Acknowledgements This work was supported by the Natural Sciences and Engineering Research Council of Canada. We thank Marg Ingleton for technical assistance and Hal Pashler and Rob McCann for comments on an earlier draft of the article. Roberto Dell'Acqua is currently supported by a postdoctoral fellowship from Foundation Fyssen.

References

Arnell, K. M., & Jolicœur, P. (in press). The attentional blink across stimulus modalities: Evidence for central processing limitations. *Journal of Experimental Psychology: Human Perception and Performance.*

Blake, R. R., & Fox, R. (in press). Visual form recognition threshold and the psychological refractory period. *Perception & Psychophysics, 5,* 46–48.

Broadbent, D. E., & Broadbent, M. H. P. (1987). From detection to identification: Response to multiple targets in rapid serial visual presentation. *Perception & Psychophysics, 42,* 105–113.

Chun, M. M., & Potter, M. C. (1995). A two-stage model for multiple target detection in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 109–127.

Coltheart, M. (1980). Iconic memory and visible persistence. *Perception & Psychophysics, 27*, 183–228.

De Jong, R., & Sweet, J. B. (1994). Preparatory strategies in overlapping-task performance. *Perception & Psychophysics, 55*, 142–151.

Duncan, J. (1980). The locus of interference in the perception of simultaneous stimuli. *Psychological Review, 87*, 272–300.

Duncan, J., Martens, S., & Ward, R. (1997). Restricted attentional capacity within but not between sensory modalities. *Nature, 387*, 808–810.

Duncan, J., Ward, R., & Shapiro, K. L. (1994). Direct measurement of attentional dwell time in human vision. *Nature, 369*, 313–315.

Giesbrecht, B. L., & Di Lollo, V. (1998). Beyond the attentional blink: Visual masking by item substitution. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 1454–1466.

Green, D. M., & Swets, J. A. (1974). *Signal detection theory and psychophysics.* Huntington, NY: Krieger.

Jolicœr, P. (in press-d). Modulation of the attentional blink by on-line response selection: Evidence from level of choice difficulty in unspeeded vs speeded $Task_1$ decision. *Journal of Experimental Psychology: Human Perception and Performance.*

Jolicœur, P. (in press-a). Modulation of the attentional blink by on-line response selection: Evidence from speeded and unspeeded $Task_1$ decisions. *Memory & Cognition.*

Jolicœur, P. (in press-b). Restricted attentional capacity between sensory modalities. *Psychonomic Bulletin & Review.*

Jolicœur, P. (in press-c). Dual-task interference and visual encoding. *Journal of Experimental Psychology: Human Perception and Performance.*

Jolicœur, P., & Dell'Acqua, R. (1998). The demonstrations of short-term consolidation. *Cognitive Psychology, 36*, 138–202.

Meyer, D.E., & Kieras, D.E. (1997). A computational theory of executive cognitive processes and human multiple-task performance: Part 2. Accounts of psychological refractory-period phenomena. *Psychological Review*, *104*, 749–791.
Pashler, H. (1989). Dissociations and dependencies between speed and accuracy: Evidence for a two-component theory of divided attention in simple tasks. *Cognitive Psychology, 21*, 469–514.
Pashler, H. (1993). Dual-task interference and elementary mental mechanisms. In D. E. Meyer & S. Kornblum (Eds.), *Attention and performance: XIV. Synergies in experimental psychology, artifical intelligence, and cognitive neuroscience* (pp. 245–264). Cambridge, MA: MIT Press.
Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin, 116*, 220–244.
Potter, M. C. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory, 2*, 509–522.
Potter, M. C. (1993). Very short-term conceptual memory. *Memory & Cognition, 21*, 156–161.
Potter, M. C., Chun, M. M., Banks, B. S., & Muckenhoupt, M. (1998). Two attentional deficits in serial target search: The visual attentional blink and an amodal task-switch deficit. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 979–992.
Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance, 18*, 849–860.
Scheerer, E. (1973). Integration, interruption and processing rate in visual backward masking. *Psychologische Forschung, 36*, 71–93.
Shibuya, H., & Bundesen, C. (1988). Visual selection from multielement displays: Measuring and modeling effects of exposure duration. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 591–600.
Shulman, H., & Hsieh, V. (1995). *The attention blink in mixed modality streams*. Paper presented at the 36th Annual Meeting of the Psychonomic Society, Los Angeles, CA.
Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied, 74*, 1–29.
Van Selst, M., & Jolicœur, P. (1994). A solution to the effect of sample size on outlier elimination. *Quarterly Journal of Experimental Psychology, 47A*, 631–650.