

BRIEF REPORTS

Restricted attentional capacity between sensory modalities

PIERRE JOLICOEUR

University of Waterloo, Waterloo, Ontario, Canada

The purpose of the experiments was to demonstrate an attentional-blink (AB) effect in a visual second task following the processing of a simple auditory signal. Subjects monitored a stream of letters presented at the middle of a computer screen using rapid serial visual presentation for the presence of a visual target (an X or a Y). In experimental trials, the visual target followed a pure tone that required an immediate speeded-choice response. When the tone had to be processed, accuracy in the visual-encoding task suffered a marked and prolonged deficit that was timelocked to the onset of the tone. When the tone could be ignored or when no tone was presented, no deficit was observed in the visual task. The results demonstrate a cross-modal AB effect produced by a simple two-choice auditory discrimination task. The results are consistent with the view that at least part of the AB effect has a central, amodal, postperceptual locus.

Dual-task interference is often found when two or more stimuli presented in the same sensory modality must be processed concurrently (e.g., Duncan, 1980) or in rapid succession (e.g., Duncan, Ward, & Shapiro, 1994; Raymond, Shapiro, & Arnell, 1992). The latter interference was referred to as an attentional blink (AB) by Raymond et al. (1992). Such interference can be explained by postulating restrictions in processing capacity at several possible levels in the information-processing stream. Capacity limitations could occur in within-modality sensory and perceptual systems or in later, more cognitive stages of processing (such as encoding information into short-term memory). Considerable evidence, both physiological and psychological, suggests that early sensory and early perceptual mechanisms have large processing capacity and that capacity limitations occur at later stages, where information must be integrated and stored.

In this view, early within-modality mechanisms have high capacity and can process several stimuli in parallel. Later mechanisms may be restricted to serial processing. Surprisingly, however, recent evidence using dual-task experiments (Duncan, Martens, & Ward, 1997) suggests that there may be more interference from the joint processing of two stimuli in the same modality than from two in different modalities, suggesting a within-modality

locus of attentional capacity limitation. The main goal of this article was to demonstrate that capacity limitations in a visual perceptual-input task can also be found when two stimuli are presented in different sensory modalities, in a paradigm that is similar to the one used to study the AB phenomenon.

Several studies have investigated the time course of attentional demands of visual encoding using the following general approach. Under dual-task conditions, both of two target stimuli presented in close temporal contiguity had to be processed. In the control condition, only one of the two stimuli required processing. Capacity limitations in processing mechanisms were revealed by a relative performance deficit in the dual-task condition relative to the single-task control condition. The time course of the attentional demands associated with processing the first stimulus were measured by changing the lag, or stimulus-onset asynchrony (SOA), between the first and the second target stimuli in the dual-task condition (e.g., Raymond et al., 1992). This technique was used to investigate whether an attention-demanding task requiring the processing of an auditory stimulus—making a simple decision about the pitch of a pure tone—would cause a deficit in the encoding of a visual target presented concurrently with the auditory task. The two experiments in this article show that making a speeded response to a pure tone caused a sustained attentional deficit in a concurrent attention-demanding visual encoding task. The results suggest that some aspect of visual encoding may require attentional mechanisms that can be occupied by the concurrent processing of a stimulus presented in a different sensory modality. Implications of the results are considered in the General Discussion.

This work was supported by the Natural Sciences and Engineering Research Council of Canada. I thank Marg Ingleton for technical assistance and Jeremy Wolfe, Hal Pashler, John Duncan, and Roddy Roediger for comments and criticisms of an earlier draft of this article. Correspondence should be sent to P. Jolicoeur, Department of Psychology, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: pjolicoe@cgl.uwaterloo.ca).

EXPERIMENT 1

In every trial, letters were shown by using rapid serial visual presentation (RSVP) at the center of a computer display. The stream contained either the letter X or the letter Y (but never both). The visual task was to report, at the end of each trial, which target letter (X or Y) had been shown. This response was made without speed pressure.

A pure tone was presented on every trial, concurrently with the RSVP stream. Control subjects were instructed to ignore the tone. Experimental subjects made a speeded response (pressing one of two buttons) depending on the pitch of the tone. Both groups performed the visual task on every trial. The control condition provided a test for the possibility that the mere presentation of a (concurrent) tone might be sufficient to cause a deficit in the visual task (e.g., Davis, 1959).

Method

Subjects. Students at the University of Waterloo with normal or corrected-to-normal vision and normal hearing participated as paid volunteers. There were 10 subjects, 19–23 years (mean 20.6), in the control condition and 9 subjects, 19–23 years of age (mean 20.6), in the experimental condition.

Stimuli. The auditory stimuli were pure tones presented for 100 msec at a frequency of either 400 or 1,200 Hz. The tones were presented by using the internal speaker in the computer. The visual stimuli were uppercase letters (1° visual angle) presented on a computer screen in white on a black background. The letters were presented in RVSP, at the same location at the center of the screen, for 100 msec each with no blank interstimulus interval. There were 6–9 letters (randomized at run time) presented prior to the letter concurrent with the tone and 9–12 letters (also randomized at run time) following the tone. The X or Y could occur, with equal probability, in any of the eight positions following the tone.

Procedure. Each trial began with two symbols at the center of the screen; these provided both fixation markers and performance feedback for the previous trial. A press of the spacebar initiated the RSVP sequence. Subjects in the experimental group responded to the tone, by pressing the > key if the tone had a low pitch or the ? key if the tone had a high pitch, as quickly as possible while keeping errors to a minimum. If the response time (RT) to the tone was greater than 1,100 msec, a message was presented asking the subject to respond more quickly to the tone. At the end of every trial (for experimental subjects after the response to the tone), a prompt asked the subject to indicate which visual target had been shown (X or Y). The X key was used to respond "X," and the C key was used to respond "Y." This response was not speeded. The subjects were tested individually in a single session that consisted of two blocks of 32 practice trials followed by five blocks of 64 experimental trials. Each block of trials contained two replications of all possible combinations of the two tone frequencies, the two visual targets, and the eight SOAs. Different random orders of the trials were used for each block and for each subject. For additional details concerning the procedure and stimuli, see Jolicoeur (1998).

Results and Discussion

Figure 1 (filled symbols, solid lines) shows the average accuracy in the visual encoding task (Task₂) depending on the interval of time (SOA) separating the tone and the visual target (X or Y) for the control and experimental groups.

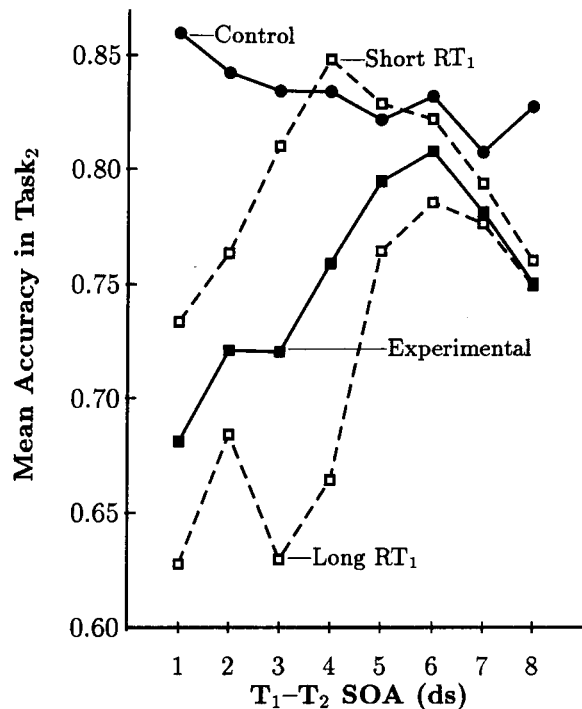


Figure 1. Results from Experiment 1. Filled symbols: Mean accuracy in the visual encoding task depending on the time interval (SOA, in deciseconds) between the auditory stimulus and the visual target and on whether the auditory signal was processed (experimental condition) or ignored (control condition). Open symbols: Mean accuracy in the visual encoding task for the experimental condition depending on response time in the auditory task.

Mean accuracy scores in the visual task were computed for each SOA for each subject and submitted to a mixed-model analysis of variance (ANOVA) in which groups (control vs. experimental) was a between-subjects factor and SOA was a within-subjects factor. Only trials with a correct response in the tone task were used in these analyses. Accuracy was higher, on average, in the control group than in the experimental group [$F(1,17) = 7.84$, $MS_e = .031600$, $p < .012$], suggesting that processing the tone engaged attentional mechanisms that were required to perform the visual-encoding task. What is more interesting is that the deficit in the experimental group was time locked with the onset of the tone. Accuracy in the experimental group was lower immediately after presentation of the tone and recovered as the SOA was increased. In contrast, in the control group, accuracy was generally higher than in the experimental condition and declined slightly as SOA was increased, although this effect was not significant in a separate ANOVA [$F(7,63) = .48$, $MS_e = .0048$, $p > .84$]. The pattern of results produced a statistical interaction between group and SOA [$F(7,119) = 2.69$, $MS_e = .005400$, $p < .013$]. The main effect of SOA was not significant [$F(7,119) = .81$, $MS_e = .005400$, $p > .58$]. A separate analysis of the results from the ex-

perimental group revealed a significant effect of SOA [$F(7,56) = 2.71$, $MS_e = .006146$, $p < .017$].

The results from the control group show that the mere presentation of a salient auditory event concurrently with the visual-encoding task did not cause an attentional deficit. However, when a response that required the processing of the tone had to be performed, a substantial deficit in the concurrent visual-encoding task was observed. This difference in the patterns of results as a function of SOA, across the experimental and control groups, shows that performing the tone task interfered with some aspect of the processing required to perform the visual task.

Figure 1 (open squares, dashed lines) also shows accuracy in the visual task for the experimental condition for trials conditionalized on the speed of the response in the auditory task (Task₁). The trials for each subject, at each SOA, were divided into two bins on the basis of the median response time in that cell. The mean accuracy was then determined for each bin for each subject. Accuracy was lower, overall, when a long response time was observed in the auditory task than when a short response time was observed [$F(1,8) = 11.76$, $MS_e = .022177$, $p < .009$]. Furthermore, there was an interaction between speed of response and SOA [$F(7,56) = 2.24$, $MS_e = .009292$, $p < .045$]. The accuracy functions converged at longer SOAs, as would be expected given that the response had already been made on most trials (the mean response time was 366 msec for faster responses and 564 msec for slower responses).¹

One interpretation of the correlation between RT₁ and accuracy in Task₂, which is supported by converging experimental (i.e., not correlational) evidence (Jolicoeur, in press-a), is that the duration of processing in the auditory task had a direct consequence for concurrent processing in the visual task. In this view, the results suggest that a shorter duration of processing in the auditory task releases attentional processes required for the visual task earlier, thereby leading to better performance in the visual task.

An ANOVA, with SOA as a within-subjects factor, was also performed on the response times in the tone task for the experimental group. The RTs were first screened for outliers using a variant of the procedure recommended by Van Selst and Jolicoeur (1994), which resulted in the exclusion of 69 of 2,704 correct trials (2.6%).² The mean response time was 449 msec, and did not vary significantly across SOA [$F(7,56) = .70$, $MS_e = 530.604$, $p > .66$].

Accuracy scores in the tone task (for subjects in the experimental condition) were also computed for each subject at each SOA, and submitted to an ANOVA with SOA as a within-subjects factor. The mean accuracy was 94% and did not vary as a function of SOA [$F(7,56) = 1.13$, $MS_e = .001611$, $p > .35$].

EXPERIMENT 2

In Experiment 1, control trials were performed under single-task instructions and the experimental and control

functions did not converge completely at longer SOAs. This lack of convergence suggests that some of the difference between the two conditions could be attributed to general costs associated with performing two tasks. In Experiment 2, experimental and control trials were performed by each subject under more equivalent conditions of dual-task load. Within each block of trials, a tone was presented on half of the trials. When the tone was presented (experimental condition), a speeded response was required. The control condition was created by having trials in which the tone was not presented. On these trials, the subject still had to remain prepared to perform the speeded auditory task (because the tone could be presented at any time), but the auditory-tone task was not engaged (because the tone was not presented).

Method

Subjects. Sixteen students at the University of Waterloo, 18–25 years of age (mean 20.6), participated as paid volunteers. All reported having normal or corrected-to-normal vision and normal hearing.

Stimuli and Procedure. The stimuli were the same as in Experiment 1, but the procedure differed in the following way. A within-subjects design was used in which an equal number of experimental and control trials were intermixed at random within each block of trials. When the tone was presented (experimental trials), the subject was required to make an immediate speeded response. When the tone was not presented (control trials), the subject pressed the spacebar, after termination of the RSVP stream, without speed pressure. An unspeeded choice response was made at the end of both types of trials to indicate whether X or Y had occurred in the RSVP stream.

Each subject performed 2 blocks of 32 practice trials followed by 10 blocks of 64 experimental trials containing all combinations of two tone frequencies, two visual targets, and the eight SOAs.

On tone-absent (control) trials, one of the letter positions in the RSVP stream was nevertheless designated as the position for the tone (were one to have been presented). In this way, a lag or SOA could be defined, relative to the hypothetical position of the tone, even for tone-absent trials. In the case of control trials, therefore, SOA codes and provides a control for potential effects associated with the absolute temporal position of the visual target within the RSVP stream.

Results and Discussion

The results are shown in Figure 2.

Again, processing the auditory signal in the simple speeded choice task produced a marked, lengthy, and time-locked deficit in the visual-encoding task. In contrast, accuracy in the control condition was roughly constant across the corresponding positions in the RSVP stream. This pattern of results produced a highly significant statistical interaction between SOA and trial type (control vs. experimental) [$F(7,105) = 5.51$, $MS_e = .004393$, $p < .0001$]. The overall mean difference between conditions was also highly significant [$F(1,15) = 39.94$, $MS_e = .003440$, $p < .0001$], as was the main effect of SOA [$F(7,105) = 8.71$, $MS_e = .005119$, $p < .0001$].

Figure 2 (open squares, dashed lines) also shows that, as in Experiment 1, a larger deficit in the visual task was associated with a longer duration of processing in the au-

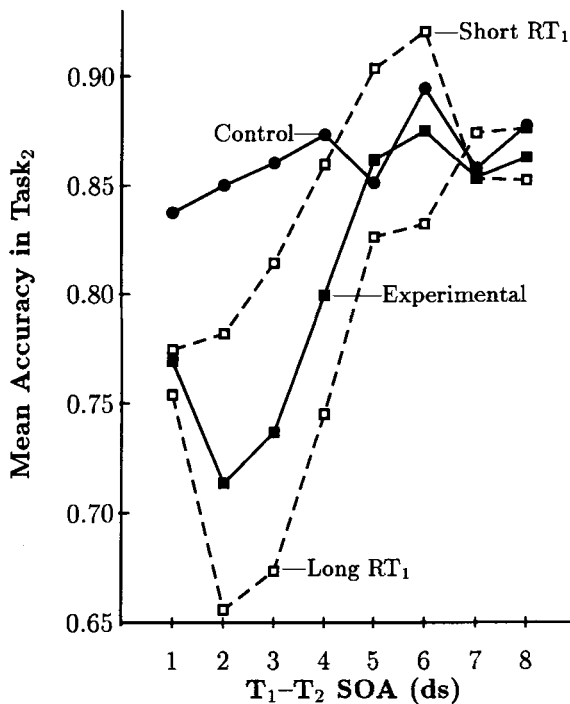


Figure 2. Results from Experiment 2. Filled symbols: Mean accuracy in the visual encoding task depending on the time interval (SOA, in deciseconds) between the auditory stimulus and the visual target, when the auditory signal was presented (experimental condition) or when the auditory signal was not presented (control condition). Open symbols: Mean accuracy in the visual-encoding task for the experimental condition, depending on the response time in the auditory task.

ditary task. The lower accuracy in Task₂ (the visual task) for trials with a longer RT₁ (RT in the tone task) than for trials with a shorter RT₁ was statistically significant [$F(1,15) = 30.50$, $MS_e = .008939$, $p < .0001$], as was the interaction of this factor with SOA [$F(7,105) = 3.65$, $MS_e = .009330$, $p < .0015$]. The mean response time in the auditory task was 420 msec for response times shorter than the median and 583 msec for response times longer than the median.

The mean response time to the tone was 502 msec, and did not vary significantly across SOA [$F(7,105) = .81$, $MS_e = 435.677$, $p > .57$] (2.0% of the trials were rejected as outliers). The mean accuracy on tone-present trials was 93% and did not vary as a function of SOA [$F(7,105) = 1.16$, $MS_e = .001558$, $p > .33$].

GENERAL DISCUSSION

The results in both experiments have shown that the processing of a simple auditory signal in an attention-demanding speeded-choice RT task caused a substantial and time-locked interference in a concurrent visual-encoding task that did not require a speeded response (Figures 1 and 2). The main goal of this article was to

demonstrate such interference in the context of a paradigm similar to that used to study the AB phenomenon (e.g., Raymond et al., 1992). The results leave no doubt that it is possible to produce an AB-like effect in a visual second task by the concurrent processing of a stimulus presented in a different sensory modality.

One prominent account of the AB phenomenon is that the AB effect reflects capacity limitations in access to, and retrieval from, visual short-term memory (VSTM; Shapiro, Raymond, & Arnell, 1994). This account cannot easily explain the present findings because it seems unlikely that the presentation of a pure tone would give rise to a representation in VSTM. The VSTM account of the AB effect could be augmented by supposing that there was more than one locus, or more than one type, of interference contributing to the observed performance deficit studied under the rubric of the AB phenomenon.

Three other accounts of the AB phenomenon have been proposed (Chun & Potter, 1995; Duncan et al., 1994; Jolicoeur, 1998). In these accounts, visual input is identified to a deep level without significant capacity limitation (late selection). However, the representations activated at this level (called level 1 in Duncan, 1980) cannot be used for further processing (e.g., storage in memory or response selection) without an additional capacity-demanding transformation to what Duncan (1980) called level 2. Chun and Potter (1995) and Duncan and his colleagues (Duncan et al., 1997; Duncan et al., 1994) have developed models of the AB effect based on the notion that the fundamental cause of the AB effect is in a capacity-demanding transformation of representations from level 1 to level 2. In their original conception, both of these proposals assumed that AB-like interference would be produced only by within-modality stimulation. Jolicoeur (1998, in press-a) and Jolicoeur and Dell'Acqua (1998) proposed that the mechanisms that transfer information to STM are susceptible to interference from central processing involved in operations such as response selection, mental rotation, retrieval from LTM, or other "central processes" (see Jolicoeur, 1998). The expression "central processes" is meant to be interpreted in the same sense as that used by Pashler (1994). These are processes that appear to play a key role in the psychological refractory period (PRP) paradigm. If a component of the AB effect is associated with such central interference, then it should be possible to demonstrate interference associated with the processing of a stimulus presented in a different modality, as was found in this article. The amount of information loss in the visual task would depend on multiple factors, such as the rate of information loss caused by the masking effects of other stimuli in the presentation sequence and by the degree and duration (see Figures 1 and 2, open squares) of interference caused by the central processing engaged by a concurrent competing task (Jolicoeur, 1998, in press-a, in press-b).

The theoretical frameworks developed by De Jong and Sweet (1994) and by Rogers and Monsell (1995) are also

useful in understanding the present results. According to De Jong and Sweet, dual-task interference depends both on direct interference and on what they called "task preparation." They argued (and provided supporting evidence) that concurrent tasks could be prepared to varying degrees, and that the efficiency of processing in a given task depended on the degree to which it was prepared. Suppose that both the tone task and the visual task required attention-demanding coordination of cognitive operations. Suppose further that maintaining sets of processes in a prepared state is, itself, an ability that is capacity limited. In this view, the performance of the speeded tone task could have prevented the subjects from preparing optimally for the visual-encoding task. This lowered state of preparation could account for the main-effect difference between the control and experimental groups in Experiment 1. Preparation alone cannot account for the AB effect, however, because performance remained high in the control condition of Experiment 2 (tone not presented) despite the fact that subjects had to be prepared to perform the tone task. Perhaps, once the tone task was completed, "preparation" could be shifted to the visual task. This notion is similar to the concept of "task switching" (e.g., Rogers & Monsell, 1995). If so, we would also expect to see a recovery of performance in the visual task as the lag (or SOA) between the tone and the visual target was increased. Furthermore, an earlier termination of the operations required to perform the tone task would also allow preparation to shift to the visual task earlier, which could explain the more rapid recovery of performance in the visual-encoding task on trials with a short RT_1 relative to trials with a long RT_1 . Potter, Chun, Banks, and Muckenhoupt (1998) have argued that there are two distinct types of AB effects: one due to direct within-modality interference and another due to costs of task-switching. They argued that cross-modal AB effects were due entirely to task-switching costs. In my view, what is most critical is that interference associated with preparation and task-switch costs would not occur if the visual encoding task did not require central mechanisms (assuming we can rule out peripheral loci of interference; see Jolicoeur, in press-b).

The present results could be seen as conflicting with those of Pashler (1989). He found only negligible interference effects on report of visually encoded information when a masked visual display followed a simple tone task at a short SOA. There are many differences between the present experiments and those of Pashler (1989). Pashler used a more complex visual task involving simultaneously displayed multielement visual arrays and longer exposure durations. It is still not clear which of these or other differences may be responsible for the differences in results, and more work will be required to elucidate this apparent conflict. It is possible that the RSVP procedure used in this article provided a more sensitive test of interference effects on visual encoding.

Furthermore, De Jong and Sweet (1994) showed that differential task preparation could have large effects on the results of the Pashler (1989) paradigm. They argued (and provided evidence) that Pashler's subjects were relatively less prepared for the tone task and more prepared for the visual task and that this pattern of differential task preparation tended to produce smaller interference effects in the visual encoding task.

The obvious cross-modal AB effects in Experiments 1 and 2 were also quite different from the null cross-modal effects reported by Duncan et al. (1997). There were several differences in the experimental procedures used in the two studies, however, which could explain this apparent discrepancy. One major difference was the rate of stimulus presentation within sensory modality. In the present experiments, a new visual stimulus was presented every 100 msec. In the Duncan et al. auditory-visual condition, a new visual stimulus was presented once every 250 msec. It is possible that this difference in presentation rate could provide the basis for an explanation of the apparent discrepancy in experimental results. Other methodological differences could also have contributed to the different results. It is unlikely that the use of an immediate speeded response in Task₁ in the present work as opposed to a deferred unsped response in Duncan et al. (1997) was a determining factor, because Arnell and Jolicoeur (in press) and Potter et al. (1998) both found cross-modal AB using deferred and unsped Task₁ responses. Additional work will be required to pinpoint the boundary conditions of the present effects and those in Duncan et al. (1997) and Pashler (1989). What is clear, however, is that it is possible to produce large AB-like effects in a visual-encoding task involving an unsped deferred response as a result of processing a very simple auditory stimulus. Models of human information processing will need to account for these interference effects. My hope is that this article will provide incentives to pursue these issues in future work.

REFERENCES

- ARNELL, K. M., & JOLICOEUR, P. (in press). The attentional blink across stimulus modalities: Evidence for central processing limitations. *Journal of Experimental Psychology: Human Perception & Performance*.
- 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 & Performance*, **21**, 109-127.
- DAVIS, R. (1959). The role of "attention" in the psychological refractory period. *Quarterly Journal of Experimental Psychology*, **11**, 211-220.
- 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.
- JOLICOEUR, P. (1998). Modulation of the attentional blink by on-line re-

- response selection: Evidence from speeded and unspeeded Task₁ decisions. *Memory & Cognition*, **26**, 1014-1032.
- JOLICOEUR, P. (in press-a). Concurrent response selection demands modulate the attentional blink. *Journal of Experimental Psychology: Human Perception & Performance*.
- JOLICOEUR, P. (in press-b). Dual-task interference and visual encoding. *Journal of Experimental Psychology: Human Perception & Performance*.
- JOLICOEUR, P., & DELL'ACQUA, R. (1998). The demonstration of short-term consolidation. *Cognitive Psychology*, **36**, 138-202.
- 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. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin*, **2**, 220-244.
- POTTER, M. C., CHUN, M. M., BANKS, B. S., & MUCKENHOUP, 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, & 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 & Performance*, **18**, 849-860.
- ROGERS, R. D., & MONSELL, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, **124**, 207-231.
- SHAPIRO, K. L., RAYMOND, J. E., & ARNELL, K. M. (1994). Attention to visual pattern information produces the attentional blink in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception & Performance*, **20**, 357-371.
- VAN SELST, M., & JOLICOEUR, P. (1994). A solution to the effect of sample size on outlier elimination. *Quarterly Journal of Experimental Psychology*, **47A**, 631-650.

NOTES

1. Performance in both tasks likely improved with practice, which could have produced an association between good performance levels across tasks. This possibility was tested by splitting the data into five blocks of trials and then repeating the analysis with blocks as a factor. The association was present within each block of trials, ruling out common improvement with practice as a third-variable account of the correlation. The association was also present within each block of Experiment 2.

2. Van Selst and Jolicoeur (1994) advocated computing the mean and standard deviation of the observations in a cell after the temporary exclusion of the largest observation in the cell, for distributions with positive skew. In the present analyses the algorithm temporarily excluded the score that was furthest from the mean, which, as expected, was the largest observation in most cases (all but five iterations of the algorithm). See Jolicoeur (1998) for additional details.

(Manuscript received November 24, 1997;
revision accepted for publication March 13, 1998.)