Revisiting within-modality and cross-modality attentional blinks: Effects of target–distractor similarity

KAREN M. ARNELL
Brock University, St. Catharines, Ontario, Canada

and

RYAN JENKINS
North Dakota State University, Fargo, North Dakota

When two masked targets (T1 and T2) require attention and are presented within half a second of each other, the report accuracy for T2 is reduced, relative to when the two targets are presented farther apart in time. This effect is known as the attentional blink (AB). Potter, Chun, Banks, and Muckenhoupt (1998) argued that all AB-like effects observed when at least one of the targets was presented outside of the visual modality did not represent true instances of the AB, but instead were artifacts of task-set switching. However, in the Potter et al. experiments the presence or absence of task-set switching opportunities was confounded with the T2 task, as well as the alphanumeric class of T2 with respect to the distractors. In the present experiment, we examine the influence of T1 alphanumeric class, T2 alphanumeric class, and switching operations in a fully crossed design that unconfounds these factors. In contrast to the conclusions of Potter et al., the present results suggest that the T2 alphanumeric class can account for the pattern of ABs observed across conditions, without necessarily implicating a separate switch cost. The implications for theoretical models of the AB and the debate over the validity of cross-modal ABs are discussed.

When a second target stimulus (T2), is presented before the processing of a first target (T1) has been completed, interference is often observed. In the attentional blink (AB) paradigm, both targets are masked, and an unspeeded response to each target is required. When the targets are presented within approximately half a second of each other, report accuracy for T2 is particularly poor (Broadbent & Broadbent, 1987; Raymond, Shapiro, & Arnell, 1992). In contrast, T2 report is not impaired when T2 is presented more than half a second after T1, or when T1 is unattended (Raymond et al., 1992). Almost all theoretical models of the AB suggest that T2 performance suffers while limited-capacity attentional resources are occupied with the processing of T1. However, models differ as to whether the required, but limited, attentional processes are thought to be modality specific or modality independent. For example, modality-specific theories have suggested that processing limitations underlying the AB reside in modality-specific stores (Duncan, Martens, & Ward, 1997; Soto-Faraco & Spence,

2002) or are uniquely visual in nature (Potter, Chun, Banks, & Muckenhoupt, 1998). However, modality-independent theories have suggested limits on stimulus consolidation in amodal working memory stores (e.g., Arnell & Jolicœur, 1999; Arnell & Larson, 2002; Jolicœur, 1998, 1999; Jolicœur & Dell'Acqua, 1998).

Several findings are consistent with a later, postperceptual locus of the effect but do not necessitate a central (amodal) processing limitation. For example, behavioral (Shapiro, Driver, Ward, & Sorensen, 1997) and electrophysiological (Vogel, Luck, & Shapiro, 1998) studies have demonstrated that T2 is semantically activated even when it cannot be reported accurately during the AB. Second, ABs have been observed when T1 was unmasked and required a speeded response (as in the psychological refractory period [PRP] paradigm), even when the two targets were presented in different stimulus modalities (Arnell & Duncan, 2002; Jolicœur, 1998, 1999). Under these conditions, the size of the AB also varies with the response selection requirements of T1 (Arnell & Duncan, 2002; Jolicœur, 1998, 1999), suggesting that the processing limitation underlying the AB may interact with response selection requirements.

Findings suggesting that the AB can be observed when the two critical targets are presented outside of the visual modality, and/or when T1 and T2 are presented in different stimulus modalities (e.g., Arnell & Jolicœur, 1999; Arnell & Larson, 2002; Duncan et al., 1997; Hillstrom,

Shapiro, & Spence, 2002; Shulman & Hsieh, 1995; Soto-Faraco et al., 2002) seemingly provide compelling evidence that amodal processing limitations underlie the AB. However, these experiments have produced a pattern of mixed results and have not provided the convincing evidence needed for a central processing limitation on stimulus consolidation. When both targets were presented auditorily, many studies have observed reliable auditory ABs (Arnell & Jolicœur, 1999; Arnell & Larson, 2002; Arnell, Trangsrud, Jones, & Larson, 2004; Duncan et al., 1997; Mondor, 1998; Shulman & Hsieh, 1995; Soto-Faraco & Spence, 2002). An AB has also been observed when both of the targets were tactile in nature (Hillstrom et al., 2002). However, Potter et al. (1998) reported no auditory AB unless the T1 and T2 required different tasks. Cross-modal ABs have been reported in several studies in which the two targets have been presented in different stimulus modalities (Arnell, Helion, Hurdelbrink, & Pasieka, 2004; Arnell & Jolicœur, 1999; Arnell & Larson, 2002; Shulman & Hsieh, 1995, for auditory–visual target combinations; and Soto-Faraco et al., 2002, for visual–tactile target combinations), but have been absent in other prominent works (Duncan et al., 1997; Potter et al., 1998; Soto-Faraco & Spence, 2002). For example, although Duncan et al. (1997) observed reliable ABs in their within-modality visual condition and within-modality auditory condition, no AB was observed in the auditory–visual cross-modal condition, prompting the authors to suggest that the AB reflects limitations on stimulus encoding in within-modality stores.

Potter et al. (1998) observed cross-modal ABs only in situations where T1 and T2 required different tasks. In light of these findings, Potter et al. suggested that the auditory and cross-modal ABs observed in all previous studies were not actually instances of the AB, but instead resulted from the time it took to switch task-set from T1 to T2. In many of the above auditory and cross-modal AB experiments, participants were asked to perform one task for T1 and another task for T2. For example, in the Arnell and Jolicœur (1999) experiments, participants in the visual T1–auditory T2 condition were asked to report the identity of the lone digit (T1) and report whether a letter X was present or absent among the subsequent letter fillers. Under these conditions, participants could have begun each trial looking for the digit and then started to listen for the X only after the digit had been detected. As discussed by Potter et al., such a strategy would produce a switch in task-set between T1 and T2. Reconfiguring one's task-set takes time, and during this period of reconfiguration immediately after T1, one may be increasingly likely to miss T2. Indeed, researchers (e.g., Allport, Styles, & Hsieh, 1994; Rogers & Monsell, 1995) have found substantial task-set reconfiguration costs outside the AB paradigm, with predictable task switches such as those in Arnell and Jolicœur. Therefore, if task-set reconfiguration takes approximately 500 msec, such switching could produce a pattern of results that resembles the AB.

Discussion of task-set switching has also focused on the issue of +1 sparing. The term *+1 sparing* refers to a pattern where T2 accuracy is higher when T2 comes immediately after T1, than when T2 comes one or two items later (see Visser, Bischof, & Di Lollo, 1999, for a review of +1 sparing). Although for many ABs, T2 accuracy is lowest at the shortest T1–T2 intervals, it is also common for the AB pattern to show +1 sparing. The presence of an AB with +1 sparing in a given experiment provides evidence against switching explanations for the results, since the detrimental effects of switching should be maximal at the shortest intervals. Furthermore, Potter et al. (1998) argued that the *absence* of +1 sparing could be taken as evidence that switching effects were operating instead of, or in addition to, the AB. Thus, the absence of +1 sparing in the results of Arnell and Jolicœur (1999) suggested to Potter et al. that auditory and cross-modal AB effects were simply artifacts of task-set switching.

Soto-Faraco and Spence (2002) suggested that in addition to task-set switching, spatial location switching from T1 to T2 was also a possible reason for the cross-modal AB patterns. In all of the auditory and cross-modal AB studies discussed above (with the exception of Soto-Faraco & Spence, 2002) auditory and visual stimuli appeared to come from different spatial locations (auditory stimuli presented over headphones or from a different speaker location, visual stimuli presented on a computer screen). It takes time to move attention from one location to another location (e.g., Posner, 1980). If a switch from the T1 location to the likely location of T2 was initiated immediately after T1, and T2 could not be successfully identified during the location switch, such switching could produce a pattern of results that resembles the AB. Using a design where visual and auditory targets were presented in the same spatial location and task-set switching was removed, Soto-Faraco and Spence observed reliable unimodal ABs for visual and auditory targets, but no cross-modal AB. In fact, T2 accuracy improved at short T1–T2 separations in one of the cross-modal conditions, leading the authors to suggest that previous cross-modal ABs were artifacts of task-set and/or spatial location switching from T1 to T2.

Despite the plausibility of switching arguments, several pieces of evidence suggest that the presence or absence of switching cannot account for the constellation of auditory and cross-modal AB findings. For example, Arnell, Trangsrud, et al. (2004, Experiment 5) observed an auditory and a cross-modal AB when the T1 and T2 tasks were the same, and when T1 and T2 identities were independent. Soto-Faraco and Spence (2002) also observed an auditory AB when T1 and T2 shared the same task and T2 identity was independent of T1 identity. Arnell and Larson (2002) observed reliable auditory and cross-modal ABs (auditory T1–visual T2, with +1 sparing) when using a design in which target identity and modality were completely random, target to target. Under these conditions, participants could not prepare to do one task and then reconfigure their attentional set to another task.

Thus, these results provide evidence that top-down task-set reconfiguration is not necessary to produce auditory and cross-modal AB patterns. Moreover, null cross-modal AB effects have been reported even under experimental conditions that promoted task and location switching (Arnell & Larson, 2002; Duncan et al., 1997; Shulman & Hsieh, 1995). For example, Arnell and Larson (2002) and Shulman and Hsieh (1995) observed no AB when T1 was visual and T2 was auditory, even though T1 and T2 were presented in different locations and required different tasks. Duncan et al. (1997) observed no cross-modal ABs despite the fact that the target locations differed, and target identities were not independent. Potter et al. (1998) observed no cross-modal ABs even though their auditory and visual targets were presented in different locations. These results provide evidence that the presence of task-set reconfiguration or location switching is not sufficient to find auditory and cross-modal AB patterns. Therefore, although switching operations can modulate the pattern of AB effects, their apparent presence appears neither necessary nor sufficient to produce AB-like patterns.

Furthermore, the studies with null cross-modal effects and those stressing the importance of switching raise some important questions. For example, the presentation conditions varied widely for the within-modality and cross-modality conditions of Duncan et al. (1997), making it difficult to know whether the null cross-modal effects resulted from these presentation differences or from the nature of the AB. Indeed, recent results from Arnell, Transgrud, et al. (2004) suggest that the pattern of results in Duncan et al. may be attributable to these presentation differences across modalities. Soto-Faraco and Spence (2002), who observed no cross-modal ABs under conditions where there was no task or location switching, had no control condition to test whether cross-modal ABs could be observed under similar conditions when task and location switching were present. Without such a control experiment, it is impossible to know whether the null cross-modal ABs would have still been observed with similar stimuli and tasks in the presence of task and location switching.

The study by Potter et al. (1998) has received particular attention because it contains an experiment where ABs are found for all four auditory–visual target combinations under conditions that promote task-set switching, as well as experiments where no auditory or cross-modal ABs are observed under conditions where no task-set switching is required. As such, it seemed reasonable to attribute the AB-like effects to the presence of task-set switching. However, the presence or absence of task-set switching in these experiments was confounded with the nature of the T2 stimulus. In Experiments 1–5 of Potter et al., the visual and auditory distractors that made up the rapid serial visual presentation (RSVP) and rapid auditory presentation (RAP) streams were digits. In Experiments 1, 2, 3, and 5, both T1 and T2 were letters, and the task was to identify the letter target(s). Experiment 6 contained a purely auditory condition where the distractors were letters and both T1 and T2 were digits, and the

task was to identify the digits. Thus, in each of these experiments T1 and T2 were from the same alphanumeric class as each other (i.e., both digits or both letters) and were presented among distractors of a different alphanumeric class (letters or digits). As discussed by Potter et al., such conditions should discourage task-set switching. In these experiments, Potter et al. observed no robust auditory or cross-modal ABs, despite the presence of robust visual ABs.

However, in Experiment 4 of Potter et al. (1998), T1 was a digit, T2 was the letter x, and the distractors were letters. Participants were instructed to identify the lone digit and to report whether a subsequent x was present or absent. This was the same task combination used by Arnell and Jolicœur (1999), and as discussed by Potter et al., such a combination promotes the use of task-set switching. Potter et al. observed reliable ABs in all four auditory–visual modality combinations in Experiment 4. Because Experiment 4 was the only experiment that promoted task-set switching and was also the only experiment to produce auditory and cross-modal AB patterns, Potter et al. concluded that the AB-like patterns were not instances of the AB, but were instead artifacts of task-set switching. However, Experiment 4 of the Potter et al. study was also the only experiment where T2 was from the same alphanumeric class as the distractors (an X among other letters). Experiment 4 was also the only experiment where T2 required a present/absent decision to a fully specified T2 (i.e., "Detect the X"). All other experiments required the identification of T2 that was part of a target set (i.e., "What digit was T2?").

The featural and/or categorical similarity of targets and distractors has been shown to modulate the size of the AB when stimuli are presented in the visual modality. Chun and Potter (1995) observed a larger AB when digits followed T1 and T2 letters than when keyboard symbols followed T1 and T2 letters and concluded that the distractor set influenced the threshold criterion for detecting a target and initiating its consolidation. Isaak, Shapiro, and Martin (1999) also found that distractors conceptually similar to T2 were more likely to be confused with T2, and produced a larger AB. Maki, Couture, Frigen, and Lien (1997) observed attenuated ABs for word targets presented among consonant strings or false-font distractors, relative to word and nonword distractors. Raymond, Shapiro, and Arnell (1995) observed a larger AB when the item trailing a T1 letter was also a letter compared with when it was a dot pattern of similar spatial frequency. Thus, it is unclear whether the auditory and cross-modal ABs in Experiment 4 of Potter et al. (1998) resulted from task-set switching as they claim, or because T2 was to be detected from the same category as the distractors. Their results cannot distinguish between these possibilities, given that the occurrence of each is perfectly confounded.

The Present Experiment

An important part of understanding the AB is determining whether central, modality-independent process-

ing limitations contribute to the AB, and if so, what conditions modulate their influence. Currently, it is unclear whether or not the categorical similarity of targets to the distractors can moderate cross-modal and auditory AB results, and whether an effect of T2 similarity could account for the findings of Potter et al. (1998). In the present experiment, we sought to determine whether the alphanumeric category of T1 and T2, with respect to the distractors, could account for the pattern of results observed by Potter et al. In our experiment, participants were asked to identify T1 and T2 on each trial. The distractors were always letters, and T1 could be a letter or a digit, as could T2. Each of the four combinations of letter/digit (LL, DD, DL, LD) was presented in each of four modality combinations (VV, AA, AV, VA)[1] so that the effects of letter/digit combination could be examined overall, as well as separately for each of the modality combinations.

When T2 was a letter, it was from the same alphanumeric class of the distractors, and when T2 was a digit, it was from a different alphanumeric class than the distractors. If the AB is more robust when T2 is from the same alphanumeric class as the distractors, and T2 class is responsible for the pattern of results in Potter et al. (1998), the AB should be more likely to be observed when T2 is a letter (LL and DL conditions), compared with when T2 is a digit (DD and LD conditions). The 2 × 2 nature of the design not only allows us to look at the unconfounded effects of T2 class, but also the unconfounded effects of T1 class (LL and LD vs. DD and DL) and the unconfounded effects of whether T1 and T2 are from the same class (LL, DD) or from a different class (LD, DL). Target modality (visual or auditory), alphanumeric class (letter or digit), and identity varied randomly and independently target to target. At the start of each trial, the participants did not know the class, modality, or identity of T1 or T2. Even after the presentation of T1, they had gained no information as to the class, modality, or identity of T2.

There are two possible types of switch costs at work in the AB paradigm. If participants know in advance of a trial that they will do one task for T1 and a different task for T2, they may purposely engage in a top-down switching strategy to tune their attention for the first task only, and then reconfigure their task-set for the T2 task after T1 has been presented. The conditions of the present experiment should discourage such *preparatory* task-set switching, given that the participant cannot usefully prepare for the modality or class of any target. However, although the present design discourages preparatory task switching, it is always possible that at least some participants would engage in this strategy anyway by simply deciding to guess on each trial, or keeping a constant switching strategy that would pay off on a minority of trials. Second, even if the participant does not prepare to switch tasks, modality and/or class readiness for T2 may be influenced by the nature of T1—that is, bottom-up switching in which the participants do not purposefully change their task-set from T1 to T2, but the system be-

comes biased in a bottom-up manner toward the class and/or modality of T1, such that doing the same task again becomes momentarily easier than the switching of mental processing needed to perform a different task. For example, having just seen a digit target, one may be better able to process another visual digit for a short period of time, and this may come at the expense of auditory targets and/or letter targets that follow soon after. Thus, although the design of the present study discourages preparatory task-set switching, stimulus-driven switch costs are still possible. If any form of task switching is present in the experiment, switch costs can be observed by comparing T2 performance when T1 and T2 are from different classes (DL, LD) with T2 performance when T1 and T2 are from the same class (LL, DD), and by comparing T2 performance in the within-modality conditions (VV, AA) with T2 performance in the cross-modality conditions (AV, VA).

Given that preparatory switching is minimized here, the present experiment has less potential for switch costs than Potter et al.'s (1998) Experiment 4 where both preparatory and stimulus-driven switch costs were possible. However, if any switch costs are present here, larger ABs should be found for the DL and LD conditions compared with those observed in LL and DD conditions. Therefore, switching theories (Potter et al., 1998) would predict pseudo-AB patterns in all modality combinations (the result of switch costs) for DL and LD conditions, but no AB patterns for LL and DD conditions outside the visual modality (i.e., no AB-like patterns in any of the auditory and cross-modal conditions if no switch costs are present with this design). Location switching theories predict AB patterns in VA and AV conditions for all letter/digit combinations (because all letter/digit combinations require the same switch in location from T1 to T2) with cross-modal targets. However, we predict that LL and DL conditions will have larger ABs than DD and LD conditions in all modality combinations because T2 will be from the same alphanumeric class as the distractors when it is a letter, but from a different class when it is a distractor.

METHOD

Participants

The participants included 40 male (*n* = 10) and female (*n* = 30) North Dakota State University undergraduate students ranging in age from 18 to 29 years, with a mean age of 23.2 years. They participated individually in a single session lasting approximately 1.5 h. The participants received credit toward a psychology course in exchange for their time. All participants reported normal (or corrected-to-normal) visual acuity and normal hearing.

Design

The design of the experiment was a 2 (T1 modality) × 2 (T2 modality) × 2 (T1 class) × 2 (T2 class) × 4 (T1–T2 interval) factorial. Each target was presented either visually or auditorily, creating four modality combinations (VV, AA, AV, VA). Each target could be either a letter or a digit, creating four letter/digit combinations (LL, DD, DL, LD). The onsets of the two targets were separated by 80, 160, 400, or 560 msec. Each of the variable levels varied randomly

within participants with the constraint that each possible combination of the factors occurred once every 64 trials. The modality and class of one target were random and independent of the other target within a trial, and across trials. Thus, the participants never knew the modality combination or letter/digit combination for the upcoming trial, nor could they know this for the second target once the first target had been presented. Each participant performed 768 trials in a single session with a break after the first 512 trials.

Apparatus and Stimuli

The experiment was controlled and timed with PsyScope software (Cohen, MacWhinney, Flatt, & Provost, 1993) and a Power Macintosh G3 computer, with a 17-in. Macintosh color monitor. The participants initiated trials and made responses using the keyboard. Auditory stimuli were presented through Sony MDR-V100 headphones connected to the computer via Harman Kardon HK195 speakers.[2] No sounds were presented directly through the computer speaker or the external speakers.

Visual distractor stimuli included all of the letters of the alphabet, with the exception of the letters J, W, and X. The letters J and X were used as the letter targets. The numbers 2 and 4 served as the digit targets. There were no digit distractors. All visual stimuli were capitalized and presented in 24-point Geneva font. At this size, the letters subtended approximately 1.0º of visual angle in height and width at an unfixed binocular viewing distance of approximately 40 cm. Visual stimuli were presented in black with RSVP, where each stimulus is presented one at a time in the same spatial location. Each visual stimulus was presented in the center of a uniform gray screen for 13 msec, followed by a 67-msec blank interstimulus interval (ISI) during which only the gray background was visible.

The auditory distractor stimuli included all of the letters of the alphabet with the exception of the letters J, W, and X. As with the visual stimuli, the letters J and X were the auditory letter targets. The numbers 2 and 4 also served as the auditory digit targets. As with the visual stream, there were no auditory digit distractors. The auditory stimuli were recordings of spoken letters and digits presented in compressed speech. The auditory recordings were the same as those used by Arnell and Jolicœur (1999), except that the duration of each letter recording was further compressed from 90 msec to 80 msec in the present study. The shorter stimuli were also those used by Arnell and Larson (2002). To create the initial stimuli, digit vocal recordings of a male voice were collected with an Apple microphone and a Power Macintosh AV computer. Recordings were done with 16 bits of resolution for amplitude at a sampling rate of 47 kHz with the aid of SoundEdit 16 software. In the present experiment, auditory stimuli were presented one at a time in RAP streams. Each auditory stimulus was presented to the right ear for 80 msec, with no blank ISI. All auditory stimuli were presented using 16 bits of amplitude resolution during the experiment.

Pilot testing indicated that it was too difficult for the participants to monitor concurrent visual and auditory streams for the four targets unless the targets were distinguished from the distractors in some manner. We wanted to choose cues that were coincident with the targets but that were not actually features of the targets. We chose to distinguish the visual targets from the distractors by adding a "*" symbol just above each visual letter and digit target (less than .50º of the top of the target). The onset and offset of the symbol were coincident with those of the visual target. No symbol appeared with the visual distractors. We chose to distinguish the auditory targets from the distractors by presenting a 530-Hz tone to the left ear for 50 msec with each auditory letter and digit target. The onset of the tone was coincident with the onset of the target letter. No tone was presented with the auditory distractors, or at any other point during the experiment.

Procedure

The participants were familiarized with the auditory targets (J, X, 2, 4) before beginning any trials. Approximately 12 practice trials preceded the experimental trials. The participants performed practice trials until they demonstrated to the experimenter that they could identify one visual target and one auditory target successfully in isolation.

The participants initiated each trial by pressing the keyboard spacebar. Each trial began with the presentation of a black fixation cross in the center of the screen for 500 msec, followed by a 500-msec blank interval before the start of the RSVP and RAP streams. The visual and auditory streams began at the same time, ran concurrently, and had the same number of stimuli. The stimulus asset asynchrony (SOA) of 80 msec for stimuli in both the visual and auditory streams produced a presentation rate of just over 12 letters/sec. Auditory and visual distractor letters were chosen randomly and independently by the computer with the constraints that distractors could not be from the target set (2, 4, J, X), and no letter could be presented twice in the same modality within a trial. When a target was presented in one modality, a random distractor was presented in the other modality. The independence across modality streams did not allow the participants to use the stimuli in one modality to assist them with their response to stimuli in the other modality. Two targets were presented on each trial. Whether the letter target was a J or an X, was chosen randomly by the computer, as was whether the number target was a 2 or a 4. The T1 and T2 identities were independent such that the identity of T1 in no way constrained the identity of T2. Eleven letters were presented in each modality before T1. T2 was presented one, two, five, or seven items after T1. Twenty-one auditory stimuli and 21 visual stimuli were presented on each trial.

Participants were informed that the modality and class (letter, digit) of the targets would be random on all trials, and therefore they should monitor both streams for any of the four possible targets at all times. They were asked to identify both targets on each trial. They were told that the "*" symbol would be presented just above the visual targets, and that an auditory tone would be presented concurrently with the auditory target. The participants were asked to identify both targets on each trial, guessing if necessary. They were told that the targets would always be from the set (2, 4, J, X) and that only these responses would be allowed. They were also informed that the same target could be presented twice on each trial (e.g., two Js). Immediately after each pair of streams, but without speed pressure, the participants were prompted by a sentence on the computer screen to press the key matching the identity of the first target. After entering a response, they were then prompted by another sentence on the screen to press the key matching the identity of the second target. Accuracy was stressed, and participants were aware that their response times were not being recorded.

RESULTS

T2 Accuracy

Figure 1 shows the mean T2 accuracy (% correct) as a function of modality combination, letter/digit combination, and T1–T2 interval. Figure 2 shows the same mean T2 accuracy as a function of letter/digit combination and T1–T2 interval, collapsed across target modality. T2 accuracy was calculated regardless of T1 accuracy. Given the difficult nature of the tasks, and the fact that there were a very large number of conditions (64), most participants had one or two conditions with very low T1 accuracy (the exact conditions varied across participants). Thus, unconditional T2 performance is presented here because the means for each participant are more precise than when T2 accuracy is conditionalized upon T1 being correct. However, all key aspects of the data are also observed when T2 accuracy was conditionalized on correct
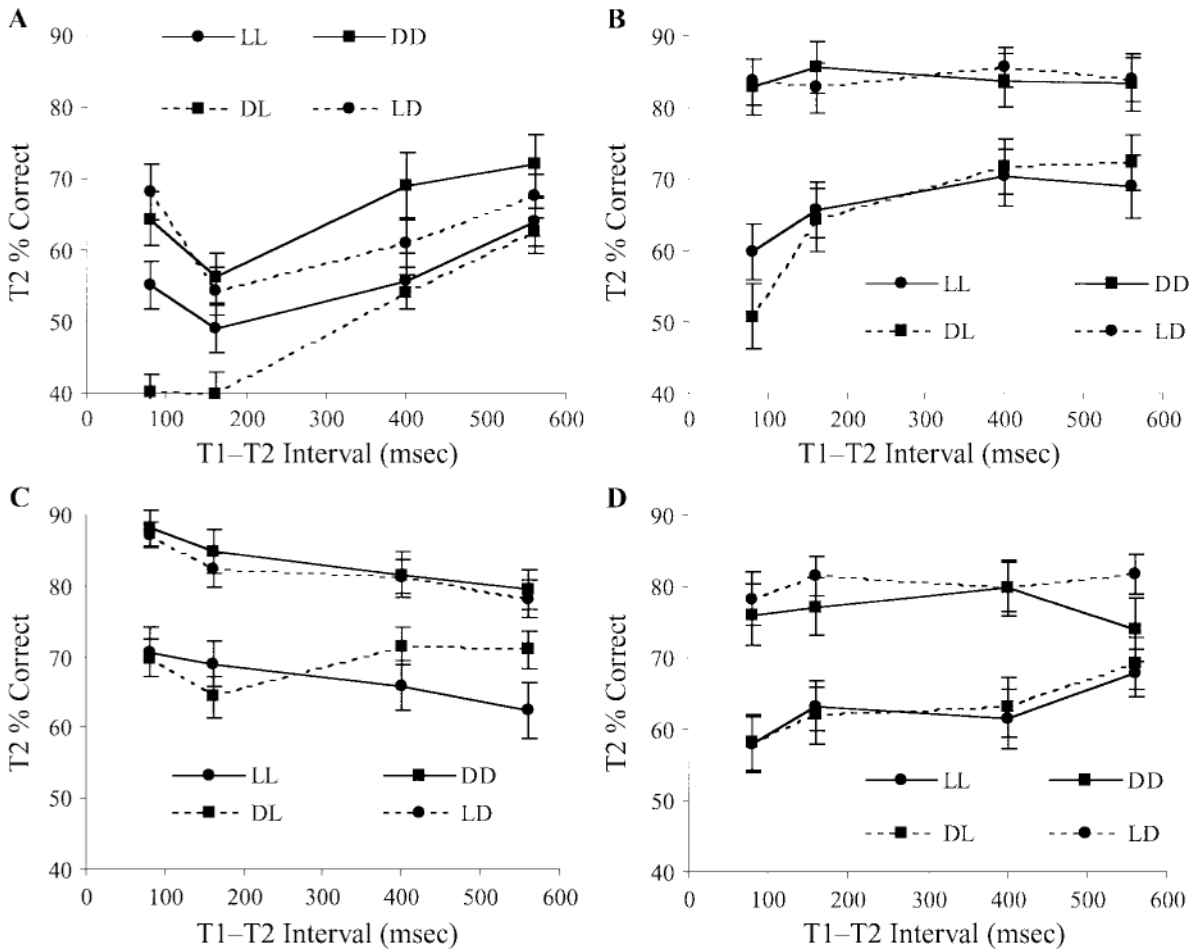
A

B

C

D

Figure 1. The group mean percentage of correct T2 responses as a function of target class, target modality, and T1–T2 interval. Panel A contains means obtained when both targets were presented visually. Panel B contains means obtained when both targets were presented auditorily. Panel C contains means obtained when T1 was presented auditorily and T2 was presented visually. Panel D contains means obtained when T1 was presented visually and T2 was presented auditorily. Error bars represent the standard error for each cell mean.

T1 report, and any discrepancies in the results with the two procedures are noted below.

All trials where T1 and T2 had the same identity (e.g., both 2), regardless of modality combination, were removed prior to the analysis to eliminate the potential for repetition blindness effects (e.g., Kanwisher, 1987; Miller & MacKay, 1994). Responses were scored as correct even if they were entered in the order opposite to their presentation (i.e., T2 response followed by T1 response) so that possible order confusions at short T1–T2 intervals would not appear to be possible ABs.

As suggested by Figure 1, all letter/digit combinations in the VV modality condition produced a reliable AB, with the DL condition producing the numerically largest AB. However, both the auditory and cross-modal conditions showed no AB pattern when T2 was a digit, despite the presence of reliable ABs in these modality conditions when T2 was a letter. Figure 2 shows that across the

modality conditions, the largest AB was produced by the DL condition, and that no AB was observed when T2 was a digit (DD or LD conditions). Below, the results are analyzed separately for each modality condition and overall. For all tests, significance was evaluated with an alpha level of .05.

T2 analyses for each modality condition. Repeated measures analyses of variance (ANOVAs) were performed on T2 accuracy rates for each modality combination (VV, AA, AV, VA) to examine the effects of T1 class, T2 class, and T1–T2 interval for the modality conditions separately. This allowed an examination of the possible impact of T1 and T2 class on T2 accuracy across the T1–T2 intervals for each of the modality conditions. When both targets were visual (VV, Figure 1A)[3] the analysis showed a main effect of T2 class [$F(1,39) = 13.97$, $MS_e = 1,509.75$ $p < .001$], a main effect of T1–T2 interval [$F(3,117) = 12.08$, $MS_e = 636.58$, $p < .001$], and a sig-
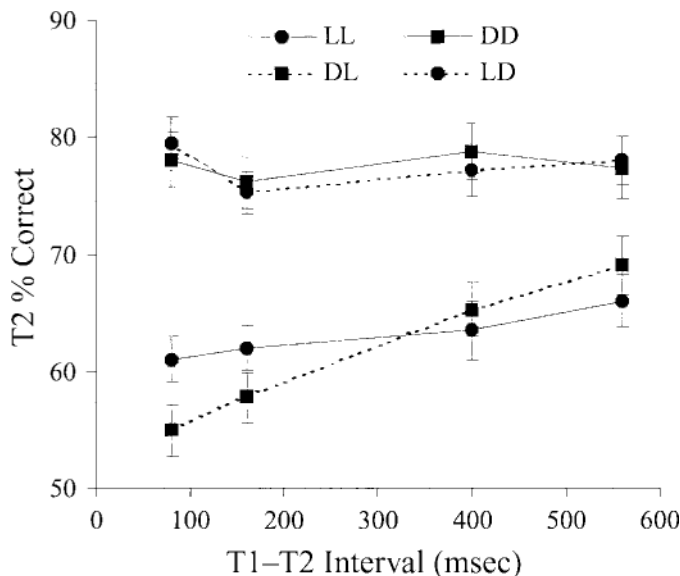
Figure 2. The group mean percentage of correct T2 responses, collapsed across modality conditions, as a function of target class and T1–T2 interval. Error bars represent the standard error for each cell mean.

nificant T2 class × T1–T2 interval interaction [$F(3,117) =$ 3.01, $MS_e = 351.23$, $p < .05$]. T2 accuracy was lower overall when T2 was a letter, and T2 accuracy deficits at short T1–T2 intervals were more pronounced for letter T2s. There was no main effect of T1 class [$F(1,39) =$ 1.76, $MS_e = 375.04$, $p > .19$], nor did T1 class enter into a three-way interaction with T2 class and T1–T2 interval ($F < 1$). However, the T1 class × T1–T2 interval interaction was significant [$F(3,117) = 5.04$, $MS_e = 251.73$, $p < .01$], with T1–T2 interval having more effect when T1 was a digit than when T1 was a letter. The T1 class × T2 class interaction was also significant [$F(1,39) = 9.66$, $MS_e = 357.47$, $p < .01$], with T2 class having a stronger effect on T2 accuracy when T1 was a digit. Planned one-way ANOVAs performed separately for each letter/digit combination showed a reliable effect of T1–T2 interval for all four of the letter/digit combinations (all $ps < .05$), with all combinations showing an AB-like pattern. Numerically, all letter/digit combinations also produced a +1 sparing pattern, with the exception of the DL condition, although the +1 sparing was not observed in the LL condition when T2 accuracy was conditionalized on T1 accuracy.

When both targets were auditory (AA, Figure 1B), the analysis showed a main effect of T2 class [$F(1,39) =$ 53.30, $MS_e = 1,031.06$, $p < .001$], a main effect of T1–T2 interval [$F(3,117) = 13.94$, $MS_e = 174.06$, $p < .001$], and a significant T2 class × T1–T2 interval interaction [$F(3,117) = 11.52$, $MS_e = 166.20$, $p < .001$]. T2 accuracy was lower overall when T2 was a letter, and a larger effect of T1–T2 interval was observed when T2 was a letter than when T2 was a digit. There was no main effect of T1 class ($F < 1$) and no interaction between T1 class and

T2 class ($F < 1$). The T1 class × T1–T2 interval interaction was not quite significant [$F(3,117) = 2.13$, $MS_e =$ 158.39, $p > .09$]. The three-way interaction with T1 class, T2 class, and T1–T2 interval was significant [$F(3,117) =$ 2.97, $MS_e = 117.71$, $p < .05$], resulting in part from the particularly large effect of T1–T2 interval in the DL condition, and absent T1–T2 interval effects in the DD and LD conditions. Planned one-way ANOVAs performed separately on each letter/digit combination showed a reliable effect of T1–T2 interval for the LL ($p < .01$) and DL ($p < .001$) combinations, both of which showed an AB pattern (with no +1 sparing), but not for the DD ($p >$ .83) and LD ($p < .47$) combinations where T2 accuracy did not vary across the T1–T2 interval.

When T1 was auditory and T2 was visual (AV, Figure 1C), a significant main effect of T2 class was observed [$F(1,39) = 47.23$, $MS_e = 614.08$, $p < .001$], where T2 accuracy was lower overall when T2 was a letter. There was a main effect of T1–T2 interval [$F(3,117) =$ 5.23, $MS_e = 200.79$, $p < .01$], but T2 accuracy was higher at shorter intervals, which is the opposite of the typical AB pattern. The T2 class × T1–T2 interval interaction was not significant [$F(3,117) = 1.67$, $MS_e = 213.94$, $p >$ .17]. There was a marginally significant main effect of T1 class [$F(1,39) = 3.23$, $MS_e = 466.92$, $p < .10$], where T2 accuracy was lower when T1 was a letter than when T1 was a digit. T1 class did not interact with T1–T2 interval [$F(3,117) = 1.76$, $MS_e = 187.56$, $p > .15$], nor with T2 class ($F < 1$). The three-way T1 class × T2 class × T1–T2 interval interaction was not quite significant [$F(3,117) = 2.23$, $MS_e = 202.71$, $p > .05$]. Planned one-way ANOVAs performed on each letter/digit combination showed a reliable *improvement* in T2 accuracy at

short T1–T2 intervals for the DD ($p < .05$) and LD ($p < .001$) conditions. The accuracy improvement at short SOAs was not reliable in the LL condition ($p > .24$). Only in the DL condition was a reliable AB-like pattern with +1 sparing observed ($p < .05$).

When T1 was visual and T2 was auditory (VA, Figure 1D), the analysis revealed a significant main effect of T2 class [$F(1,39) = 52.50$, $MS_e = 747.04$, $p < .001$], where T2 accuracy was lower overall when T2 was a letter than when T2 was a digit. There was a significant effect of T1–T2 interval overall [$F(3,117) = 4.04$, $MS_e = 209.85$, $p < .01$], where T2 accuracy was lower at shorter T1–T2 intervals. T2 class interacted with T1–T2 interval [$F(3,117) = 3.65$, $MS_e = 210.19$, $p < .05$], due to the larger effect of T1–T2 interval for T2 letters than for T2 digits. There was no main effect of T1 class [$F(1,39) = 1.71$, $MS_e = 225.28$, $p > .19$], nor did T1 class interact with T1–T2 interval ($F < 1$), or enter into a three-way interaction with T2 class and T1–T2 interval ($F < 1$). However, the T1 class × T2 class interaction was marginally significant [$F(1,39) = 3.52$, $MS_e = 199.08$, $p < .10$], with T2 class having a stronger effect on T2 accuracy when T1 was a letter. Planned one-way ANOVAs were performed on each letter/digit combination and showed a reliable effect of T1–T2 interval in the DL condition ($p < .01$), and a marginally significant effect in the LL condition ($p < .10$), where T2 accuracy was lower at shorter T1–T2 intervals. However, no reliable effect of T1–T2 interval was found for the DD and LD combinations ($p > .35$, and $p > .41$, respectively).

AB magnitude across the modality conditions. Additional analyses were performed, allowing a comparison of the AB size as a function of T1 and T2, class and modality. For each participant, the magnitude of the AB was calculated for each target class and modality combination. The AB size was estimated as the difference between T2 accuracy at the longest T1–T2 interval and at the interval where T2 accuracy was the most extreme, relative to the longest interval. For example, if an AB with +1 sparing was observed for a given condition, the AB size was calculated as the difference in T2 accuracy at lag 7 and lag 2. With this difference measure, the AB size is not underestimated in conditions where there is +1 sparing relative to conditions where there is no +1 sparing.[4]

The resulting AB magnitude was then used as the dependent variable in a repeated measures ANOVA with T1 modality, T2 modality, T1 class, and T2 class as factors. The analysis produced a highly significant main effect of T2 class [$F(1,39) = 33.25$, $MS_e = 413.18$, $p < .001$], where the AB was larger when T2 was a letter than when T2 was a digit. There was also a significant main effect of T1 class [$F(1,39) = 6.12$, $MS_e = 455.17$, $p < .05$], where the AB was larger when T1 was a digit than when T1 was a letter (however, this effect was not significant when T2 accuracy was made conditional upon getting T1 correct). In addition, there was a significant T1 class × T2 class interaction [$F(1,39) = 11.65$, $MS_e =$

320.57, $p < .01$], where T1 class had a significant effect ($p < .01$) when T2 was a letter (larger AB when T1 was a digit), but no effect ($F < 1$) when T2 was a digit. There was no main effect of T2 modality ($F < 1$), but there was a significant main effect of T1 modality [$F(1,39) = 34.87$, $MS_e = 421.88$, $p < .001$], and a significant T1 modality × T2 modality interaction [$F(1,39) = 31.77$, $MS_e = 747.87$, $p < .001$], which resulted in part from the particularly large AB observed in the VV condition and the particularly strong reverse effect (i.e., priming at short T1–T2 intervals) in the AV condition. Interestingly, all of the interactions, including at least one modality variable and at least one class variable, were nonsignificant (all $ps > .10$), suggesting that the effect of target class had similar effects on the AB magnitude across the modality conditions.

T1 Accuracy

Figure 3 shows mean T1 accuracy (% correct) as a function of modality combination, letter/digit combination, and T1–T2 interval. All trials were examined except those where T1 and T2 had the same identity as above. Responses were again scored as correct even if they were entered in the order opposite to their presentation (i.e., T2 response followed by the correct T1 response).

Repeated measures ANOVAs were also performed on T1 accuracy rates to examine the effects of T1 class, T2 class, and T1–T2 interval separately for each of the four modality conditions. When both targets were visual (VV, Figure 3A) the analysis showed a main effect of T1 class [$F(1,39) = 15.77$, $MS_e = 739.04$, $p < .001$], and a main effect of T1–T2 interval [$F(3,117) = 3.30$, $MS_e = 144.53$, $p < .05$]. T1 accuracy was lower overall when T1 was a letter and at short T1–T2 intervals. However, the nonsignificant T1 class × T1–T2 interval interaction [$F(3,117) = 1.61$, $MS_e = 192.43$, $p > .19$], provided evidence that the effect of the T1–T2 interval on T1 performance was the same whether T1 was a letter or a digit. The main effect of T2 class was significant [$F(1,39) = 4.40$, $MS_e = 207.44$, $p < .05$], due to slightly higher T1 accuracy when T2 was a digit. T2 class interacted with T1–T2 interval [$F(3,117) = 6.05$, $MS_e = 200.61$, $p < .001$], which reflected decreased T1 accuracy at short T1–T2 intervals when T2 was a digit, but increased T1 accuracy at short T1–T2 intervals when T2 was a letter. T2 class did not enter into a three-way interaction with T1 class and T1–T2 interval [$F(3,117) = 1.47$, $MS_e = 162.36$, $p > .22$]. However, the T1 class × T2 class interaction was significant [$F(1,39) = 40.39$, $MS_e = 266.58$, $p < .001$], with T1 class having a stronger effect when T2 was a digit. Planned one-way ANOVAs performed on T1 accuracy rates for each letter/digit combination showed a reliable effect of T1–T2 interval for three of the letter/digit combinations ($p < .001$ for DL, $p < .01$ for LD, and $p < .05$ for DD combinations). In the LD and DD conditions, T1 accuracy was reduced at short T1–T2 intervals. However, in the DL condition T1 accuracy was higher at shorter T1–T2 intervals. For the LL combination,
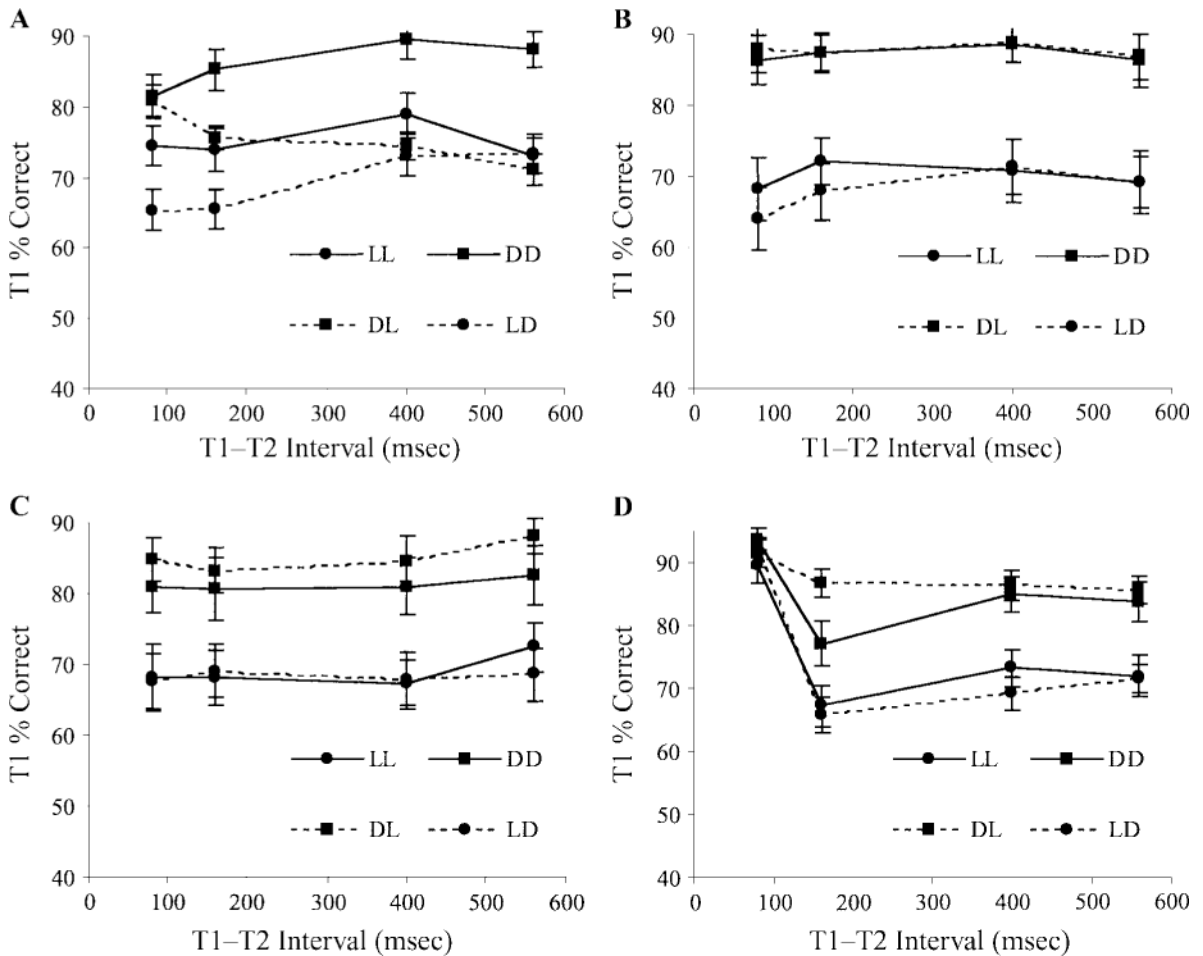
Figure 3. The group mean percentage of correct T1 responses as a function of target class, target modality, and T1–T2 interval. Panel A contains means obtained when both targets were presented visually. Panel B contains means obtained when both targets were presented auditorily. Panel C contains means obtained when T1 was presented auditorily and T2 was presented visually. Panel D contains means obtained when T1 was presented visually and T2 was presented auditorily. Error bars represent the standard error for each cell mean.

accuracy did not change reliably across T1–T2 interval ($p > .38$). These results suggest that T1 bore some of the cost in the LD and DD conditions, and that the reduced T2 accuracy in the DL condition may have resulted in part from a tendency to "overperform" on the T1 digit when the two appeared closely in time.

When T1 and T2 were both auditory (AA, Figure 3B), a significant main effect of T1 class was again observed [$F(1,39) = 37.10$, $MS_e = 1,450.97$, $p < .001$], where T1 accuracy was lower overall when T1 was a letter. No other main effects or interactions were found to be significant. Overall T1 accuracy was not influenced by T1–T2 interval [$F(3,117) = 1.91$, $MS_e = 153.42$, $p > .13$], and this was not modulated by T1 class [$F(3,117) = 1.02$, $MS_e = 122.55$, $p > .38$], T2 class ($F < 1$), or their interaction ($F < 1$). The main effect of T2 class was also not significant [$F(1,39) = 1.71$, $MS_e = 155.53$, $p > .19$], and did not interact with T1–T2 interval ($F < 1$).

When T1 was auditory and T2 was visual (AV, Figure 3C), a significant main effect of T1 class was observed [$F(1,39) = 33.59$, $MS_e = 1,015.42$, $p < .001$], where T1 accuracy was lower overall when T1 was a letter. As with the AA modality condition, there were no other main effects or interactions. Overall T1 accuracy was not influenced by T1–T2 interval [$F(3,117) = 1.93$, $MS_e = 146.40$, $p > .12$], and this was not modulated by T1 class ($F < 1$), T2 class ($F < 1$), or their interaction ($F < 1$). The main effect of T2 class was not significant [$F(1,39) = 2.80$, $MS_e = 304.54$, $p > .10$], nor was the T1 class × T2 class interaction [$F(1,39) = 1.51$, $MS_e = 275.00$, $p > .22$].

When T1 was visual and T2 was auditory (VA, Figure 3D), the analysis revealed a significant main effect of T1 class [$F(1,39) = 35.43$, $MS_e = 551.78$, $p < .001$], where T1 accuracy was lower overall when T1 was a letter than when T1 was a digit. There was a significant ef-

fect of T1–T2 interval overall [$F(3,117) = 53.85$, $MS_e = 174.96$, $p < .001$], where T1 accuracy *increased* at short T1–T2 intervals. T1 class and T1–T2 interval also produced a significant interaction [$F(3,117) = 9.80$, $MS_e = 170.69$, $p < .001$], in that the increased T1 accuracy at short T1–T2 intervals was more pronounced when T1 was a letter. There was no main effect of T2 class [$F(1,39) = 2.77$, $MS_e = 160.03$, $p > .10$], nor did T2 class interact with T1-class [$F(1,39) = 1.12$, $MS_e = 148.30$, $p > .29$], or enter into a three-way interaction with T2 class and T1–T2 interval [$F(3,117) = 1.25$, $MS_e = 157.80$ $p > .29$]. However, the T2 class $\times$ T1–T2 interval interaction approached significance [$F(3,117) = 2.65$, $MS_e = 177.14$, $p < .10$], with T1–T2 interval having a slightly stronger effect on T1 accuracy when T2 was a digit. Planned one-way ANOVAs were performed on each letter/digit combination and showed a reliable increase in T1 accuracy at short T1–T2 intervals for all four letter/digit combinations (all $ps < .05$).

DISCUSSION

Comparing the Letter/Digit Combinations

The present study examined the AB for all four possible combinations of letter and digit targets (LL, LD, DL, DD), when presented among letter distractors, and compared the T1 and T2 accuracy patterns within and across visual, auditory, and cross-modal target conditions.

Across modality conditions, it was clear that the T2 class reliably influenced the magnitude of the AB, with larger ABs when T2 was a letter than when T2 was a digit. The mean AB magnitude showed an average AB size of 11.0% for letter T2s and only 1.7% for digit T2s across modality conditions (20.4% for letter T2s and 0.6% for digit T2s when the area method was used; see Note 4). Furthermore, the larger AB for letter T2s was not restricted to just one or two of the modality conditions but was present in all four of the modality combinations (including the VV condition). Although the robustness of auditory and cross-modal ABs has previously been called into question, in the present study, when one or more targets was auditory, an AB was observed every time T2 was a letter (with the exception of the LL condition in the AV modality, and even here the priming observed with digit T2s was not significant), such that five out of six opportunities produced an AB. Furthermore, no AB was observed whenever T2 was a digit (i.e., six out of six opportunities produced no AB).

Across modality combinations, the categorical similarity of T1 to the distractors also had a small effect on the magnitude of the AB, with larger ABs when T1 was a digit than when T1 was a letter (the opposite of the T2 pattern). Across modality conditions, the mean AB magnitude was 8.4% when T1 was a digit and 4.3% when T1 was a letter (14.3% vs. 6.8%, respectively, with the area method of estimating AB size). However, closer examination makes it clear that the effect of T1 class was restricted to the case where T2 was a letter (i.e., larger AB

for DL condition than for LL condition), given that the ABs were for the most part confined to these letter/digit combinations. This interaction between T1 class and T2 class has two plausible explanations. One could argue that T1 class should have an equal effect regardless of T2 class, but that this could not be observed here given the null AB for T2 digits. Alternately, one could argue that the AB was larger in the DL condition than in the LL condition because switching contributed to the performance deficits in the DL condition, but not in the LL condition. We examine each of these in turn.

In this experiment, identification of even the first target was difficult, and T1 accuracy did not approach ceiling. T1 accuracy was higher for T1 digits (84.5%) than for T1 letters (71.4%), suggesting that attention may have been allocated to T1 more often or more fully when T1 was a digit. If this were the case, and T1 received more attentional processing when it was a digit, then it would likely have led to less attentional resources being available for T2, which would result in a larger AB. Basically, the idea here is that T1 must receive attentional processing before there can be a bottleneck on attentional processing for T2. At first, the notion of the easier T1 stimulus (the digit) demanding more attention than the harder T1 stimulus (the letter) may seem counter to previous work suggesting that the AB is larger when the T1 task has lower accuracy (Seiffert & Di Lollo, 1997). However, the majority of studies analyzed by Seiffert and Di Lollo were those where T1 clearly stood out from the distractors by virtue of a feature (color), easily capturing attention. Both views may be correct, in that an ideal T1 (to produce an AB) may readily capture attention and then require attentional processes for a long duration. If the presentation of T1 does not capture attention, no AB will be observed (so that more detectable targets are more likely to produce an AB), but once attention has been captured, the difficulty of the target discrimination determines the duration of processing and therefore predicts the size of the AB. According to this logic, the DD condition should have had a larger AB than the DL condition (just as the DL condition was larger than the LL condition), but this effect could not be observed for the DD and LD combinations because there was no AB found for these conditions.

The second possibility is that the AB was larger in the DL condition than in the LL condition because task switching contributed to the size of the AB in the DL condition. The participants may have been more prepared to process a letter T2 than a digit T2 after just detecting a T1 letter. According to this logic, the DL condition should also have had a larger AB than the DD condition, but this effect was not observed for the DD and LD combinations, because there was no AB found for these conditions. There is also some evidence in favor of this explanation, given that very little +1 sparing was observed for the DL combination relative to other letter/digit combinations, and this would be expected, given that switching would have maximal effects at the shortest target intervals.

Thus, it is possible that the larger AB observed in the DL versus the LL condition resulted from digit T1s being more effective at producing the AB, due to the time to switch from a digit T1 to a letter T2, or due to both of these. However, what is important is not whether switching may or may not have had any effect, but that there is a robust effect of T2 class that cannot be explained by switching operations, and that this effect was observed in visual, auditory, and cross-modal conditions.

Situating the Results

The larger AB observed for T2s that were from the same category as the distractors is in agreement with the results of Chun and Potter (1995), Isaak et al. (1999), and Maki et al. (1997), who showed a larger visual AB when the distractors were conceptually similar to the targets. The present results also converge with those of Taylor and Hamm (1997), who showed a larger visual AB when subjects were asked to detect a T2 "0" (pronounced "oh") among the letter distractors, than when they were told to detect the identical T2 "0" (pronounced "zero") among the letter distractors. However, the present results also provide the first evidence suggesting that the importance of T2's categorical similarity to the distractors extends to auditory and cross-modal presentations (at least for digits among letters). We do not suggest that the use of digit T2s among letter distractors can *never* lead to auditory or cross-modal AB findings. However, we do suggest that the nature of the T2 class with respect to the distractor class modulates the magnitude of the AB for all modalities, making it more or less difficult to observe AB effects, and confounding the interpretation of Potter et al.'s (1998) results.

Indeed, in the visual modality condition, although the AB was reliably larger when T2 was a letter, a significant AB was still observed for each letter/digit combination. The ABs observed in the LD and DL conditions replicate those of many previous visual AB studies where the T1 and T2 stimulus categories and/or tasks have differed (e.g., Arnell & Jolicœur, 1999, Experiments 1, 2, and 4; Potter et al., 1998, Experiment 4; Raymond et al., 1992, 1995; Shapiro, Raymond, & Arnell, 1994). The ABs observed in the visual LL and DD conditions replicate those of many previous visual ABs studies where the stimulus categories and tasks have been the same for T1 and T2 (e.g., Arnell & Larson, 2002; Chun & Potter, 1995; Potter et al., 1998, Experiments 1, 2, and 5; Soto-Faraco & Spence, 2002).

The presence of auditory and cross-modal ABs in the DL condition replicates the results of Arnell and Jolicœur (1999), and Potter et al. (1998, Experiment 4), who both observed auditory and cross-modal ABs under conditions where T1 was a digit and T2 was a letter, and distractors were letters. The presence of the DL cross-modal ABs also fit nicely with behavioral and electrophysiological data suggesting a postperceptual locus of the AB (e.g., Shapiro, Driver, et al., 1997; Vogel et al., 1998), and with recent electrophysiological evidence suggesting that the identification of T1 can slow the identifica-

tion of an unmasked T2 from the same modality (Vogel & Luck, 2002) or from a different modality (Arnell, Helion, et al., 2004).

The absence of auditory and cross-modal ABs for the DD condition also replicates Potter et al. (1998), who found no auditory AB for digit targets presented among letter distractors (Experiment 6), no auditory AB for letter targets presented among digit distractors (Experiments 1, 2, and 5), and no cross-modal ABs for letter targets presented among digit distractors (Experiments 3 and 5).[5] Soto-Faraco and Spence (2002) also observed robust improvements in T2 accuracy at short T1–T2 intervals in the AV condition when digit targets were presented among letter distractors, just as we observed here.

Potter et al. (1998) suggested that the AB-like effects from the DL experiments of Arnell and Jolicœur (1999), and Potter et al. (Experiment 4) were due to a switch of task-set from T1 to T2, instead of instances of the AB. They claimed that auditory and cross-modal AB-like effects would only be observed when a task-set switch is required and that such effects will therefore be artifactual, not true ABs. The present design discouraged preparatory task-set switching because the participants knew that the T1 and T2 modality and class were independent, and varied unpredictably from target to target. However, performance could have been influenced by bottom-up switching driven by stimulus occurrence that biased processing in favor of the T1 class (e.g., participants more prepared to process a T2 letter than a T2 digit just after processing a T1 letter). Following the logic of Potter et al., such readiness changes could underlie the apparent AB for the DL condition. If the AB-like effects in the DL condition were artifacts of switching, one would also expect AB-like effects (produced by task switching) in the LD condition of the present experiment (where T1 and T2 class also differed) and no AB-like effects in the LL condition (where T1 and T2 class were the same). However, the reverse was found. No AB was observed in the LD condition, and an AB was observed in the LL condition.

Because we have all of the 2 (T1 letter/digit) × 2 (T2 letter/digit) cells filled, we can see that, at least in the present study, the presence or absence of the auditory and cross-modal AB is not dependent on whether T1 and T2 come from the same or different classes. Instead, the presence or absence of auditory and cross-modal AB is modulated by whether T2 comes from the same or different alphanumeric class as the *distractors*. In Potter et al. (1998), the presence or absence of a task switch was always confounded with whether T2 was from the same or different class (i.e., T2 was from the same class as the distractors only in Experiment 4, where there was also the potential for a task switch). Thus, it was unclear whether task switching or T2 class was responsible, although Potter et al. chose to focus on the task-set switching possibility. Of course it is likely that task-set switching effects were more robust in Potter et al. (Experiment 4) than they were here, given that their design encouraged

such switching and the present one did not. Therefore, it is certainly possible, and indeed quite likely, that the results of Potter et al. were influenced both by T2 category and preparatory switch costs. Nonetheless, what is important from the present experiment is that: (1) the present results suggest that the alphanumeric class of T2 has large effects on the magnitude of the AB, (2) these effects are unconfounded by task switching, (3) the presence of these effects provides an alternate explanation for the results of Potter et al., and (4) the results provide evidence that auditory and cross-modal AB patterns cannot be dismissed as artifactual switch costs.

Soto-Faraco and Spence (2002) suggested that previous cross-modal AB results may not only be artifacts of task-set switching, but also artifacts of location switching, given that visual and auditory targets are routinely presented so that their perceived location differs. Indeed, when they presented visual and auditory targets so that they appeared to come from the same location, they observed visual and auditory ABs, but no cross-modal ABs. However, there was no control condition or experiment where the same visual and auditory stimuli were presented so that their perceived location differed. Under Soto-Faraco and Spence's conditions, it is difficult to tell whether the null cross-modal ABs resulted from use of the same locations, or an inability of the stimuli and/or tasks to produce cross-modal AB. In fact, their task required participants to report the digits among letter distractors—the same task that has produced null cross-modal results here. Nonetheless, it was possible that the different perceived location of visual and auditory targets could have caused artifactual location switching effects in the present experiment. Such effects, if present, should have been observed whenever a location change was present (i.e., for all VA and AV combinations, regardless of class combination). However, no AB-like effects were observed for the DD and LD combinations, or for the LL combination for the AV modality condition despite the different locations, providing evidence that location switching was not sufficient to produce AB-like effects. Instead, robust T2 facilitation was observed at short T1–T2 intervals for the AV modality, suggesting that the auditory T1 provided a cue for the presence of the visual T2. Therefore, both the absence of ABs for some of the letter/digit combinations in the cross-modal conditions, and the presence of robust facilitation at short T1–T2 intervals argue against location switching explanations for the observed cross-modal ABs.

The findings of a +1 sparing pattern and facilitation at short T1–T2 intervals in the AV modality also provide evidence against task switching or class readiness differences from T1 to T2. In the present experiment, +1 sparing was observed in all conditions where T2 was presented visually (except the DL combination in the visual modality where performance at the two shortest target intervals approximated chance levels, possibly obscuring any +1 sparing), and was never present when T2 was presented auditorily. According to Potter et al. (1998),

the presence of +1 sparing can be taken as an indicator of "true AB" that is uncontaminated by task-set switching, whereas the absence of +1 sparing suggests that the observed dual-task costs result from task-set switching in whole or in part. Although we do not agree that +1 sparing must be present in true AB, we do agree that the presence of +1 sparing strongly suggests the presence of true AB (i.e., although the absence of +1 sparing does not tell us anything about the validity of the AB, its presence does suggest true AB). The presence of +1 sparing patterns with visual T2s provides evidence that the participants were not less prepared (for whatever reason) for a visual T2 that immediately followed a T1 from a different class. If switching alphanumeric class was present, then +1 sparing effects should not have been observed for the DL and LD combinations, yet +1 sparing patterns were observed for these combinations whenever T2 was visual (with the exception of the DL combination in the visual modality).

Comparing Modality Combinations

Overall, in the present study, the VV modality had the largest AB, followed by the AA modality, and then the cross-modality conditions. This same pattern was also observed in the DL condition where the largest ABs were observed (23% difference for VV, 20% difference for AA, and approximately 10% difference for the cross-modal conditions). Observing particularly large VV ABs is not unusual (Arnell & Larson, 2002; Arnell, Trangsrud, et al., 2004; Potter et al., 1998, Experiment 4), nor is the finding of larger ABs for within-modality conditions than for cross-modality conditions (Arnell & Larson, 2002; Arnell, Trangsrud, et al., 2004; Shulman & Hsieh, 1995). However, typically the AV modality condition has produced larger ABs than the VA modality condition, which is the opposite of the pattern here. Indeed, what stands out as atypical among the results is the large T2 accuracy facilitation present at short T1–T2 intervals in the AV modality condition. Although an AB with +1 sparing was observed in the DL condition, the other letter/digit combinations from the AV modality showed higher accuracy at shorter target intervals, suggesting that the presence of an auditory T1 cued the occurrence of a visual T2 presented soon after (a similar finding was also observed by Soto-Faraco & Spence, 2002). However, participants did not enhance their visual T2 performance at the expense of the auditory T1. Auditory T1 performance in the AV condition was not influenced by T1–T2 interval. Therefore, at short target separations there was a costless priming of visual T2s by auditory T1s. This priming was also observed in the VA modality condition where visual T1 accuracy increased at short target separations, suggesting that the presence of an auditory T2 primed recently presented visual T1s (a similar finding was also observed by Arnell & Duncan, 2002). Given that this priming was as strong or stronger for the VA letter/digit combinations that did not show an AB, the finding cannot explain the VA ABs in the DL and LL conditions, and again repre-

sents a costless priming of a visual target by an auditory target. Auditory cues have been shown to successfully cue the location in space of a subsequent visual target (e.g., Spence & Driver, 1996). The present experiment, as well as those of Arnell and Duncan (2002), and Soto-Faraco and Spence (2002) were instances where T1 and T2 modality were unpredictable in order, so they may speak to the ability of an auditory event to cue a visual event under conditions of temporal uncertainty. However, although auditory targets prime visual targets, they do not prime other auditory targets (T1 accuracy was flat across target intervals, and there was no +1 sparing for T2 in the AA condition), nor do visual targets prime other visual targets. Indeed, the only example of an accuracy tradeoff was found in the VV condition where preference was given to visual digits over visual letters when the two were presented at short separations. This may have also contributed to the lack of +1 sparing in the DL condition of the visual modality.

The Nature of the Processing Limitations

The present results do not support the contention of Potter et al. (1998) that true AB (unconfounded by task-set switching) will be observed only when both targets are presented visually. Nor do they support theories suggesting that the AB results purely from confusion or processing limitations in modality-specific short-term stores (Duncan et al., 1997; Soto-Faraco & Spence, 2002). However, the present results do agree with models suggesting that there is competition among targets and distractors for prominence in working memory stores, and that targets are more likely to lose this competition if they are similar to distractors (Raymond et al., 1995; Shapiro et al., 1994). The results are also consistent with central two-stage processing bottleneck theories (Chun & Potter, 1995; Jolicœur, 1998, 1999; Jolicœur & Dell'Acqua, 1998). Bottleneck theories suggest a processing bottleneck on the entrance into a second stage of processing where the fragile and fleeting representation of a target becomes consolidated into a more permanent representation in working memory that can support conscious awareness. While T1 is being consolidated, T2 must wait until the bottleneck is freed while its representation decays or is overwritten by subsequent stimuli.

An AB is observed when T2 is followed by a mask that interrupts T2 processing (interruption masking), but not when the mask is absent, or when a mask is presented at the same time as T2 and makes T2 identification more difficult (integration masking; Giesbrecht & Di Lollo, 1998). Brehaut, Enns, and Di Lollo (1999) postulated that interruption masking of T2 is critical to producing the AB because the representation of the mask is substituted for the representation of T2 while the system is busy processing T1. Under conditions of limited attention, this "masking by object substitution" gives a bottom-up visual representation of the mask, and a top-down re-entrant representation of T2, with the latter being more likely to be discounted (Di Lollo, Enns, & Rensink, 2000). It is

possible that the substitution of T2 by trailing items is more likely to occur when T2 and the mask are similar physically, lexically, and/or categorically, but is less likely when T2 is distinct from the distractors in one or more of these regards, or is particularly salient, such as one's own name (Shapiro, Caldwell, & Sorensen, 1997). In this manner, during T1 processing, T2 letters may be more likely to be substituted for letter distractors than are T2 digits, thereby increasing the magnitude of the AB for T2 letters.

In addition to an amodal processing limitation, the larger AB observed for within-modality conditions compared with the cross-modality conditions also suggests the existence of a within-modality processing limitation. Such a pattern is consistent with several previous results, and accounts proposing that the bottlenecked processing limitations are sensitive to *both* central (amodal) and within-modality and task constraints (Arnell & Duncan, 2002; Arnell & Larson, 2002; Arnell, Transgrud, et al., 2004). As suggested by Arnell and Duncan, when targets are presented in the same modality, T2 is subject to within-modality processing limitations (possibly competition in modality-specific stores, as suggested by Shapiro et al., 1994) and central processing limitations (possibly consolidation in amodal working memory, as suggested by Jolicœur, 1998, and Jolicœur & Dell'Acqua, 1998), resulting in a larger AB than that observed when targets are presented cross-modally and T2 is subject only to central processing limitations. Because within-modality ABs result from both central and modality-specific processing limitations, they can be observed even when central processing demands are low. Because iconic memory is relatively brief and erased by a single backward pattern mask, within-modality limitations may be easy to observe in the visual modality condition where visual items may be more prone to object substitution. However, auditory processing limitations may be difficult to observe, given the relatively large capacity and long duration of echoic memory. Furthermore, because cross-modal AB is only influenced by central processing limitations, the central task demands must be sufficiently high to reveal these limitations, thus making cross-modal AB particularly difficult to observe. In the present study, use of T2 letters, which are harder to detect than T2 digits among letter distractors, may have increased central task demands, thereby prompting larger ABs for T2 letters. Thus, cross-modal AB findings may be more difficult to observe under conditions that reduce central processing demands, and T2 class with respect to the distractor may be influential in this regard.

Conclusions

The results of Potter et al. (1998) showed AB patterns for the visual, auditory, and cross-modal conditions when a task-set switch was required from T1 to T2, but an AB pattern only for the visual condition when no task-set switch was required. As such, their results appeared to provide good evidence that the auditory and cross-modal

ABs observed in Arnell and Jolicœur (1999) and in other studies (Duncan et al., 1997; Shulman & Hsieh, 1995) were the result of task-set switching. However, in the Potter et al. experiments, the presence or absence of task-set switching was confounded with the nature of the T2 task and the alphanumeric class of T2 with respect to the distractors. The present study unconfounded these factors and showed that the alphanumeric class of T2 had large effects on the magnitude of the AB observed in all modality combinations. Furthermore, the effect of T2 class could explain the pattern of results in the Potter et al. studies without requiring a task-set switch explanation of the auditory and cross-modal AB results. In contrast to Potter et al.'s conclusion that the AB is uniquely visual in nature, we conclude that there are both within-modality and amodal processing limitations. We offer an account of the present pattern of results that is consistent with other recent behavioral and electrophysiological data, by: (1) proposing the existence of both modality-specific and amodal processing limitations, (2) situating the amodal limitations in a two-stage bottleneck model (e.g., Jolicœur, 1998, 1999), and (3) suggesting that substitution of T2 by the item that trails it is more effective when T2 and the mask are from the same alphanumeric class.

REFERENCES

ALLPORT, [D.] A., STYLES, E. A., & HSIEH, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and nonconscious information processing* (pp. 421-452). Hillsdale, NJ: Erlbaum.

ARNELL, K. M., & DUNCAN, J. (2002). Shared and separate sources of dual-task cost in stimulus identification and response selection. *Cognitive Psychology*, **44**, 105-147.

ARNELL, K. M., HELION, A. M., HURDELBRINK, J., & PASIEKA, B. (2004). Dissociating sources of dual-task interference using human electrophysiology. *Psychonomic Bulletin & Review*, **11**, 77-83.

ARNELL, K. M., & JOLICŒUR, P. (1999). The attentional blink across stimulus modalities: Evidence for a central processing limitation. *Journal of Experimental Psychology: Human Perception & Performance*, **25**, 630-648.

ARNELL, K. M., & LARSON, J. M. (2002). Cross-modality attentional blinks without preparatory task-set switching. *Psychonomic Bulletin & Review*, **9**, 497-506.

ARNELL, K. M., TRANGSRUD, H. B., JONES, K., & LARSON, J. M. (2004). *Sources of interference underlying cross-modality dual-task costs*. Manuscript submitted for publication.

BREHAUT, J. C., ENNS, J. T., & DI LOLLO, V. (1999). Visual masking plays two roles in the attentional blink. *Perception & Psychophysics*, **61**, 1436-1448.

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 & Performance*, **21**, 109-127.

COHEN, J. [D.], MACWHINNEY, B., FLATT, M., & PROVOST, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments, & Computers*, **25**, 257-271.

DI LOLLO, V., ENNS, J. T., & RENSINK, R. A. (2000). Competition for consciousness among visual events: The psychophysics of reentrant visual processes. *Journal of Experimental Psychology: General*, **129**, 481-507.

DUNCAN, J., MARTENS, S., & WARD, R. (1997). Restricted attentional capacity within but not between sensory modalities. *Nature*, **387**, 808-810.

GIESBRECHT, B. L., & DI LOLLO, V. (1998). Beyond the attentional blink: Visual masking by item substitution. *Journal of Experimental Psychology: Human Perception & Performance*, **24**, 1454-1466.

HILLSTROM, A. P., SHAPIRO, K. L., & SPENCE, C. (2002). Attentional limitations in processing sequentially presented vibrotactile targets. *Perception & Psychophysics*, **64**, 1068-1082.

ISAAK, M. I., SHAPIRO, K. L., & MARTIN, J. (1999). The attentional blink reflects retrieval competition among multiple rapid serial visual presentation items: Tests of an interference model. *Journal of Experimental Psychology: Human Perception & Performance*, **25**, 1774-1792.

JOLICŒUR, P. (1998). Modulation of the attentional blink by on-line response selection: Evidence from speeded and unspeeded Task$_1$ decisions. *Memory & Cognition*, **26**, 1014-1032.

JOLICŒUR, P. (1999). Concurrent response selection demands modulate the attentional blink. *Journal of Experimental Psychology: Human Perception & Performance*, **25**, 1097-1113.

JOLICŒUR, P., & DELL'ACQUA, R. (1998). The demonstration of short-term consolidation. *Cognitive Psychology*, **36**, 138-202.

KANWISHER, N. G. (1987). Repetition blindness: Type recognition without token individuation. *Cognition*, **27**, 117-143.

MAKI, W. S., COUTURE, T., FRIGEN, K., & LIEN, D. (1997). Sources of the attentional blink during rapid serial visual presentation: Perceptual interference and retrieval competition. *Journal of Experimental Psychology: Human Perception & Performance*, **23**, 1393-1411.

MILLER, D., & MACKAY, D. G. (1994). Repetition deafness: Repeated words in computer-compressed speech are difficult to encode and recall. *Psychological Science*, **5**, 47-51.

MONDOR, T. A. (1998). A transient processing deficit following selection of an auditory target. *Psychonomic Bulletin & Review*, **5**, 305-311.

POSNER, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, **32**, 3-25.

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, & 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.

RAYMOND, J. E., SHAPIRO, K. L., & ARNELL, K. M. (1995). Similarity determines the attentional blink. *Journal of Experimental Psychology: Human Perception & Performance*, **21**, 653-662.

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

SEIFFERT, A. E., & DI LOLLO, V. (1997). Low-level masking in the attentional blink. *Journal of Experimental Psychology: Human Perception & Performance*, **23**, 1061-1073.

SHAPIRO, K. L., CALDWELL, J., & SORENSEN, R. E. (1997). Personal names and the attentional blink: A visual "cocktail party" effect. *Journal of Experimental Psychology: Human Perception & Performance*, **23**, 504-514.

SHAPIRO, K. L., DRIVER, J., WARD, R., & SORENSEN, R. E. (1997). Priming from the attentional blink: A failure to extract visual tokens but not visual types. *Psychological Science*, **8**, 95-100.

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.

SHULMAN, H., & HSIEH, V. (1995, November). *The attentional blink in mixed-modality streams*. Paper presented at the 36th Annual Meeting of the Psychonomic Society, Los Angeles.

SOTO-FARACO, S., & SPENCE, C. (2002). Modality-specific auditory and

visual temporal processing deficits. *Quarterly Journal of Experimental Psychology A*, **55**, 23-40.

Soto-Faraco, S., Spence, C., Fairbank, K., Kingstone, A., Hillstrom, A. P., & Shapiro, K. [L.] (2002). A crossmodal attentional blink between vision and touch. *Psychonomic Bulletin & Review*, **9**, 731-738.

Spence, C., & Driver, J. (1996). Audiovisual links in endogenous covert spatial attention. *Journal of Experimental Psychology: Human Perception & Performance*, **22**, 1005-1030.

Taylor, T. L., & Hamm, J. P. (1997). Category effects in temporal visual search. *Canadian Journal of Experimental Psychology*, **51**, 36-46.

Visser, T. A. W., Bischof, W. F., & Di Lollo, V. (1999). Attentional switching in spatial and non-spatial domains: Evidence from the attentional blink. *Psychological Bulletin*, **125**, 458-469.

Vogel, E. K., & Luck, S. J. (2002). Delayed working memory consolidation during the attentional blink. *Psychonomic Bulletin & Review*, **9**, 739-743.

Vogel, E. K., Luck, S. J., & Shapiro, K. L. (1998). Electrophysiological evidence for a postperceptual locus of suppression during the attentional blink. *Journal of Experimental Psychology: Human Perception & Performance*, **24**, 1656-1674.

Ward, R., Duncan, J., & Shapiro, K. L. (1996). The slow timecourse of visual attention. *Cognitive Psychology*, **30**, 79-109.

NOTES

1. For the modality condition abbreviations, the first letter refers to the modality of T1, and the second letter refers to the modality of T2. Therefore, VV refers to conditions where both targets were presented visually, AA to conditions where both targets were presented auditorily, VA to conditions where T1 was visual and T2 was auditory, and AV to conditions where T1 was auditory and T2 was visual. The LL, DD, DL, and DL abbreviations are similar, except the first letter refers to whether T1 was a letter or a digit and the second letter refers to whether T2 was a letter or a digit.

2. Potter et al. (1998), and Arnell and Jolicœur (1999) did not present visual and auditory stimuli in the same locations. Thus, different locations were also used in the present study. However, the design of the study allows us to look for effects of location switching, which, if present,

should be observed in the AV and VA modality conditions for all letter/digit combinations. This is examined in the Discussion section.

3. Performance of about 40% likely indicates null sensitivity. There are four alternatives, so chance for each buttonpress is 25%. However, because participants made both a T1 and T2 response on each trial, and their T2 response was scored as correct even if it was given as the T1 answer (to avoid order confusions masquerading as ABs), chance would be 25% + 25% = 50%. This 50% value would indicate null sensitivity too, but only T2 if participants always gave different responses for T1 and T2. However, sometimes participants gave the same response for T1 and T2 (e.g., said "J" for both), and on these trials they would have only one chance to get T2 correct. For example, if a participant always gave the same answer for T1 and T2, null sensitivity would be indicated by 25% since they effectively have one response with a 1-in-4 chance. Although we did not calculate the exact proportion of trials where our participants made the same response for T1 and T2, it seems reasonable to assume that they gave the same response for T1 and T2 on approximately 1/4 of the trials (given that T1 and T2 were the same on 1/4 of all trials and they knew this). If this is indeed the case, null sensitivity would be .75(.50) + .25(.25) = .4375 or approximately 44% correct (where .75 represents the 3/4 of trials where they had a 50% chance of being correct, and .25 represents the 1/4 of trials where they had a 25% chance of being correct).

4. When a more typical area measure was used to estimate the AB size (where T2 accuracy at the longest T1–T2 interval [560 msec] was taken to be T2 baseline accuracy, and T2 accuracy at each of the three shorter intervals [80, 160, 400 msec] was subtracted from the baseline accuracy, and then the three resulting differences were summed), the same pattern of results and significance was observed with the exception that the effect of T1 class was not significant ($p > .05$).

5. Both the digit target among letter distractors and letter targets among digit distractor experiments from Potter et al. (1998) are relevant comparisons with the DD condition here, given that all are instances where the task and set for T1 and T2 are the same, but the targets are presented among distractors from a different alphanumeric class (i.e., digit targets among letters or letter targets among digits).