Psychonomic Bulletin & Review
2002, 9 (3), 497-506

Cross-modality attentional blinks without preparatory task-set switching

KAREN M. ARNELL and JULIE M. LARSON
North Dakota State University, Fargo, North Dakota

When two masked targets (T1 and T2), both requiring attention, are presented within half a second of each other, report of the second target is poor, demonstrating an attentional blink (AB). Potter, Chun, Banks, and Muckenhoupt (1998) argued that all previous demonstrations of an AB occurring when one or more targets were presented outside the visual modality did not represent true AB but were, instead, artifactual, resulting from switching of task set. In the present experiments, T1 and T2 modalities were independent and varied randomly from trial to trial, allowing no useful preparatory task-set switching from T1 to T2. However, reliable ABs were observed when both targets were visual, when both targets were auditory, and cross-modally when T2s were visual. Furthermore, the ABs observed for cross-modality visual T2s showed the characteristic U-shaped pattern often found in AB experiments in which two visual targets are used—a pattern that should not be observed under task-set switching conditions. These results provide evidence that cross-modality AB can be found under conditions that do not allow useful preparatory task-set switching.

When two masked targets (T1 and T2), both requiring attention, are presented within half a second of each other, report of the second target is poor (Broadbent & Broadbent, 1987; Raymond, Shapiro, & Arnell, 1992). However, T2 report is good when T2 is presented more than half a second after T1 or when T1 does not require attention (Raymond et al., 1992). Raymond et al. (1992) named this pattern of results the *attentional blink* (AB). AB experiments require participants to make unspeeded responses after the presentation of all stimuli has ended. The fact that there is no need for on-line response selection has helped to convince many researchers that the AB results from relatively early processing limitations that occur before central (amodal) processing bottlenecks. In early AB experiments, only visual stimuli were used as targets and distractors, and these were often presented using the rapid serial visual presentation (RSVP) technique in which letters or words appeared rapidly one at a time in the same spatial location. Owing to the use of visual stimuli and the assumption that the AB resulted from relatively early processing limitations, the visual nature of the stimuli and their processing became incorporated into various theories of the AB (e.g., Duncan, Ward, & Shapiro, 1994; Raymond et al., 1992; Raymond, Shapiro, & Arnell, 1995; Shapiro, Raymond, & Arnell, 1994; Ward, Duncan, & Shapiro, 1996). For example, in the attentional dwell time theory of Duncan et al. (1994) and Ward et al., the AB is taken as a measure of the time course of visual attention. Shapiro et al. (1994) and Raymond et al. (1995) explained the AB as the result of confusion in an overcrowded visual short-term memory.

Arnell and Jolicœur (1999) tested visual theories of the AB by having participants attend to concurrent but independent visual RSVP letter streams and to compressed speech rapid auditory presentation (RAP) streams. The participants were asked to report the identity of a T1 number and the presence or absence of a subsequent X. T1 modality (visual or auditory) was fully crossed with T2 modality (visual or auditory), producing two within-modality conditions (both targets visual, both targets auditory) and two cross-modality conditions (T1 visual and T2 auditory, T1 auditory and T2 visual). One quarter of the participants received each modality combination. Arnell and Jolicœur found a reliable AB in each of the four modality combinations. The ABs were larger for the visual T2 conditions than for the auditory T2 conditions, but the AB sizes in the within-modality conditions were roughly equal to the AB sizes in the cross-modality conditions. Arnell and Jolicœur concluded that the AB is not a uniquely visual phenomenon. They argued that the robust ABs observed in the cross-modality conditions provided good evidence that central (amodal) processing limitations were responsible for the AB. Specifically, limitations on stimulus consolidation in working memory were proposed.

Using a variety of presentation conditions, other researchers have also reported an AB when one or both targets were presented outside the visual modality. Shulman and Hsieh (1995) reported an AB when both targets were auditory and for the auditory-T1–visual-T2 combination.

Portions of this work were presented at the November 1999 Annual Meeting of the Psychonomic Society in Los Angeles. This project was funded by an EPSCoR start-up grant provided to the first author. Correspondence concerning this work should be addressed to K. M. Arnell, who is now at Brock University, Department of Psychology, St. Catharines, ON L2S 3A1, Canada (e-mail: karnell@brocku.ca).

Duncan, Martens, and Ward (1997) also reported an AB when both targets were auditory but did not find an AB when one target was auditory and the other target was visual. Mondor (1998) found an auditory AB, using tones as targets and distractors. Soto-Faraco et al. (in press) reported ABs across modalities when one target was visual and the other was tactile. Jolicœur (1999) and Arnell and Duncan (2002) have demonstrated a variant of cross-modality AB in which a speeded response indicating the pitch of an unmasked auditory target impaired the subsequent identification of a masked visual character.

In many of the above experiments (Arnell & Jolicœur, 1999; Jolicœur, 1999; Shulman & Hsieh, 1995), participants performed one type of task for T1 and a different task for T2. Also, the participants knew the order of the tasks at the outset of the trials. In other experiments (Duncan et al., 1997; Mondor, 1998), participants performed the same task for T1 and T2, but the presence or identity of T1 could have changed the participants' expectations for the identity of T2. For example, Duncan et al. (1997) required participants to identify two target words embedded in nonword distractors. One of the targets was either *cot* or *cod*, and the other was either *nab* or *nap*. Since any of the four words could be T1, the participants had to prepare for all four words at the start of each trial. However, if T1 was *cot* or *cod*, then T2 could only be *nab* or *nap*; this change in possible targets could have led the participants to reconfigure their task set.

Indeed, in any experiment in which presentation conditions require participants to prepare for one task or stimulus and then reconfigure their readiness for a different task or stimulus, preparatory task-set switching may confound the interpretation of results. For example, if participants are told to listen for an auditory number (T1) and then to look for a subsequent visual X (T2), then after T1, the participants may reconfigure their attentional set away from the auditory modality and numbers and toward the visual modality and the letter X. This reconfiguration would likely take some amount of time. Thus, if the resulting T2 performance was impaired at short T1–T2 stimulus onset asynchronies (SOAs), the results may represent an AB or may, instead, indicate reduced attentional allocation during task-set reconfiguration. Outside the AB paradigm, researchers have found task-set reconfiguration costs even with predictable task switches (e.g., Allport, Styles, & Hsieh, 1994; Rogers & Monsell, 1995).

Potter, Chun, Banks, and Muckenhoupt (1998) have recently suggested that predictable task-set switch costs may underlie some or all of the auditory and cross-modality AB findings in the literature. These authors replicated Arnell and Jolicœur (1999), in which participants searched letter streams to identify a T1 digit and the presence or absence of a T2 letter X. Just as did Arnell and Jolicœur, Potter et al. found auditory and cross-modality ABs. However, Potter et al. then modified the T1 task by making T1 a letter and requiring the participants to identify both T1 and T2 letters embedded in distractor digits. Because the participants no longer had to reconfigure their preparatory task set between T1 and T2, changing the experiment in this way removed any task-set switching possibilities. A robust AB was found only when both targets were presented visually, but no AB was found in the auditory or cross-modality conditions. Potter et al. concluded that "true" AB can be observed only when both targets are visual and that auditory and cross-modal AB patterns reported in the literature were artifacts of task-set switching.

In the present experiments, we looked for an AB pattern in visual, auditory, and cross-modality conditions under presentation conditions that did not support preparatory task-set switching. Modality combinations, tasks, and target identities were presented randomly and independently from trial to trial and from target to target. Visual targets were distinguished by being blue letters among black letter distractors, and the participants were asked to report their identity. Auditory targets were letters distinguished by being presented along with a tone, and the participants were asked to report their identity. All four modality combinations (both visual, both auditory, visual then auditory, auditory then visual), target identities, and T1–T2 SOAs were presented randomly within each block of trials. The participants were told to expect this random presentation of modalities and target identities. Thus, the participants had to begin each trial by being prepared for both a visual and an auditory target. Once T1 had been presented, the participant would have gained no information about the modality or the task for T2 and would have to remain prepared for either a visual or an auditory target. Furthermore, because T1 and T2 letter identities were independent, the identity of T1 in no way constrained the identity of T2.

If visual, auditory, and cross-modality ABs could be observed with the present design that did not support preparatory task-set switching, this would provide strong evidence that true ABs can be found outside the visual modality and that auditory or cross-modality AB is not merely an artifact of preparatory task-set reconfiguration. If an AB were to be found in the visual condition, but not in the auditory or the cross-modal conditions, this would support Potter et al.'s (1998) claims regarding preparatory task-set switching and their conclusion that the AB is uniquely visual in nature.

The present design did not support preparatory task-set switching, in that there was no meaningful basis for task/modality switching. In all likelihood, switching would reduce the participants' performance, not enhance it. However, the design did not *prevent* the participants from engaging in a preparatory switching strategy, possibly resulting in the participants being ready for one task and modality at the expense of the other. Also, although the visual and the auditory tasks both required identification of a letter target, the features that specified the targets were different in the two modalities (blue color in the visual modality and concurrent tone in the auditory modality). In this sense, the participants could be said to have performed different T1 and T2 tasks in the cross-modality conditions. If task or target readiness is modulated by the occurrence of a previous stimulus, rather than expectancy, T1 might have shifted the state of readiness in favor of its own task (i.e., attention might have been shifted exogenously to favor the T1 task even if the participant did not engage in preparatory switching). Readiness models are characterized by *complementary* shifts in the efficiency of two tasks: As readiness for one task lessens and, therefore, the task is performed less efficiently, readiness for the other should increase, and thus it should be performed better. For example, a participant may begin each trial equally prepared to report the identity of the visually presented blue target letter or to report the identity of the auditory letter presented concurrently with the tone. However, suppose that T1 is a blue visual target. This may momentarily bias the participant's readiness so that he or she is more prepared for another blue visual target and less prepared for an auditory target and tone. The key signature of this exogenously driven switching is that there is a fixed amount of readiness. Therefore, if readiness is biased in favor of one task/modality (e.g., visual target identification), performance will show benefits for that task/modality (e.g., visual target identification), while correspondingly producing performance costs for the other task/modality (auditory target identification). However, a pattern for which costs (benefits) are observed for one task/modality, with no concomitant benefits (costs) for the other modality, would signal the absence of complementary shifts of task preparation. The present design not only reduced the motivation to switch tasks, but also allowed an examination of complementarity (costs for one task corresponding with benefits for the other task) in the AB paradigm, thus providing a direct test for the presence or absence of switching, whether preparatory or stimulus driven in nature.

EXPERIMENT 1

Method

Participants. Twelve participants (6 females) participated individually in a 1-h session. All were North Dakota State University undergraduate students participating for credit in an introductory psychology course. All the participants reported normal (or corrected-to-normal) visual acuity and normal hearing.

Design. The design was a 2 (T1 modality) × 2 (T2 modality) × 5 (SOA) factorial. T1 modality (visual or auditory), T2 modality (visual or auditory), and T1–T2 SOA (83, 250, 416, 583, and 750 msec) were all within-subjects variables that varied randomly, with the constraint that each possible combination of these factors occurred equally often every 40 trials. Each participant performed 400 experimental trials in one session.

Stimuli and Apparatus. Visual stimuli included all of the letters of the alphabet except K and L. The visual letters were capitalized and presented in 24-point Geneva font. At this size, the letters subtended approximately 1.0° of visual angle in height and width at an approximate binocular viewing distance of 40 cm. Visual stimuli were presented one at a time in an RSVP stream in the center of a uniform gray screen. Each letter was presented for 16.7 msec, followed by a 66.8-msec blank interstimulus interval (ISI), during which only the gray background was visible. All of the visual characters were black, except the visual target letters, which were dark blue.

The auditory stimuli were spoken letters presented in compressed speech. Auditory stimuli included all of the letters of the alphabet except for W. Auditory recordings were the same as those used by Arnell and Jolicœur (1999), with the exception that, in the present experiment, the duration of each letter recording was compressed from 90 to 80 msec. To create the initial auditory stimuli, digital vocal recordings of the male voice were collected using an Apple microphone and a Power Macintosh AV computer. Recordings were done using 16 bits of resolution for amplitude at a sampling rate of 47 kHz, with SoundEdit 16 software. Auditory stimuli were presented one at a time in RAP. Each auditory letter was presented to the right ear for 80 msec, followed by a 3.5-msec blank ISI.

The auditory tone was 530 Hz in pitch and 50 msec in duration. It was presented to the left ear at the same time as the auditory target letter (K or L). Tone and target letter onsets were coincident. K and L were never used as distractors in the auditory stream (i.e., they were never presented without the tone). It is extremely difficult for participants to concurrently monitor the auditory and the visual streams for targets. The visual targets were presented in blue, and the auditory targets were presented with a concurrent tone, to raise target performance above chance levels. Auditory letters and tones were presented through Sony MDR-V100 headphones connected to the computer via Harmon/Kardon HK195 speakers. No sound was presented directly through the computer speaker or the external speakers. All the auditory letters were presented by using 16 bits of amplitude resolution during the experiment.

The experiments were 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 their responses with the keyboard.

Procedure. The participants were familiarized with the target sounds (K and L) before attempting any trials. Also, approximately 12 practice trials preceded the experimental trials. The participants continued to perform practice trials until they demonstrated that they could successfully perform one visual and one auditory target in isolation. The participants initiated each trial by pressing the space bar on the keyboard. 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.

Auditory and visual streams began at the same time and ran concurrently for all the participants. The SOA was 83.5 msec for the stimuli in both the auditory and the visual streams, producing a presentation rate of almost 12.0 letters/sec. Auditory and visual letters were chosen independently by the computer, with the constraint that no letter was presented twice in the same modality within a trial. The independence across streams did not allow the participants to use the stimuli in one modality to assist them with their responses to stimuli in the other modality. Visual distractor letters were presented in black, and visual target letters were presented in dark blue. A blue target could be any letter except K or L. Auditory distractor letters were presented without a concurrent tone. An auditory target was the letter K or L presented concurrently with the tone. The number of letters presented before T1 (4, 6, 8, 10, or 12) was chosen randomly for each trial, with the constraint that auditory and visual streams should contain the same number of stimuli. Eight letters always succeeded T2, regardless of T1–T2 SOA, thus equating post-T2 masking at all T1–T2 SOAs.

Two targets were presented on each trial. On one quarter of the trials both targets were visual, on one quarter of the trials both targets were auditory, on one quarter of the trials T1 was visual and T2 was auditory, and on one quarter of the trials T1 was auditory and T2 was visual. Modality combination varied randomly from trial to trial, so that the participants never knew the combination for the upcoming trial. Furthermore, T1 and T2 modalities and target identities were independent, ensuring that even after T1 had been presented, the participant did not know the modality or identity of T2.

The participants were informed that the modality combination of the targets would be random on all the trials and, as such, they should monitor both streams concurrently for any possible targets. They were told that visual targets would be presented in blue and that these target letters should be identified. They were further informed that auditory targets would be presented concurrently with a tone and that they should decide whether the auditory letter presented with the tone was a K or an L. Immediately after each stream, but without speed pressure, the participants were prompted by a sentence presented 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 the participants were aware that their response times were not being recorded.

Results

Responses were scored as correct even if they were entered in the order opposite to presentation (i.e., a T2 response and then a T1 response). Any trials in which T1 and T2 had the same identity were removed prior to analysis to eliminate the potential for the repetition blindness effect to confound the results (e.g., Kanwisher, 1987; Miller & MacKay, 1994). Because of the two-alternative forced-choice (2AFC) nature of the auditory task and the independence of T1 and T2, half of the within-modality auditory trials were subject to removal. All performance scores for T1 and T2 were calculated independently of whether or not the response to the other target was correct. However, all key aspects of the data were also found when T2 performance was made conditional upon a correct response for T1. Mean target accuracy (percent correct) is plotted for T1 and T2 in Figures 1A–1D as a function of T1and T2 modality combination and T1–T2 SOA. Negative SOAs reflect T1 performance plotted as a function of the SOA backward from T2, and positive SOAs reflect T2 performance as a function of SOA forward from T1. Chance performance on the visual task equaled approximately 4% correct, and chance performance on the auditory task equaled 50%.

The mean target accuracies were submitted to an analysis of variance (ANOVA) with modality (visual or auditory), target modality relationship (T1 and T2 modality crossed or within-modality), target number (T1 or T2), and T1–T2 SOA as within-subjects factors. The analysis revealed that all main effects and interactions were significant (all *p*s $< .05$), including the four-way interaction [$F(4,44) = 5.16, p < .01$].

Given the significant four-way interaction, the effects of target number (T1 or T2) and T1–T2 SOA on response accuracy were analyzed separately for within- and cross-modality conditions both with visual targets and with auditory targets. In the cross-modality conditions, visual T1 performance from visual-T1–auditory-T2 trials was analyzed with visual T2 performance from auditory-T1–visual-T2 trials (i.e., the data in Figure 1C). Auditory T1 performance from auditory-T1–visual-T2 trials was analyzed with auditory T2 performance from visual-T1–auditory-T2 trials (i.e., the data in Figure 1D).

The within-modality visual condition produced a significant main effect of SOA, a significant overall difference in target accuracy for T1 and T2, and a significant SOA × target number interaction (all *p*s $< .001$). When the effect of SOA was analyzed separately for T1 and T2, using one-way ANOVAs, a significant effect of SOA was found for T2 [$F(4,44) = 18.63, p < .001$], but not for T1 ($F < 1$). As is evidenced in Figure 1A (note the different scale), despite the lack of preparatory task-set switching, the within-modality visual condition produced a large and robust AB with the characteristic U-shaped function that is often found in the AB paradigm (e.g., Chun & Potter, 1995; Raymond et al., 1992). Indeed, a paired sample *t* test showed T2 accuracy to be significantly higher at the 83-msec SOA than at the 249-msec SOA [$t(11) = 4.72$, $p < .001$]. The significant AB in this within-modality visual condition replicates Chun and Potter's (1995) demonstration of an AB under conditions in which no task-set switching should occur.

The within-modality auditory condition produced a significant main effect of SOA [$F(4,44) = 5.84, p < .001$], no significant main effect of target number ($F < 1$), and no significant SOA × target number interaction ($F < 1$). Figure 1B suggests that accuracy for both T1 and T2 was impaired at the shortest SOA. Analyses indicated a significant effect of SOA for T2 alone [$F(4,44) = 3.82$, $p < .01$], but no significant effect of SOA for T1 alone [$F(4,44) = 1.00, p > .41$]. Thus, whereas T2 responses showed reduced accuracy at short SOAs, T1 did as well (although not statistically), possibly suggesting the contribution of perceptual interference.

In the cross-modality visual condition, there was a significant main effect of SOA [$F(4,44) = 5.57, p < .001$], a significant effect of target number [$F(1,11) = 28.85, p < .001$], and a significant interaction of SOA and target number [$F(4,44) = 5.86, p < .001$]. Analyses revealed a significant effect of SOA for T2 [$F(4,44) = 4.50, p < .01$] and a significant effect of SOA for T1 [$F(4,44) = 7.00$, $p < .001$]. The AB pattern found in the statistics can be readily observed in Figure 1C. Although the AB in the cross-modality visual condition does not appear to be nearly as large as the AB observed in the within-modality visual condition, it is robust and was observed for all but 2 of the participants. Furthermore, T2 accuracy shows the U-shaped pattern often characteristic of the AB, as in the within-modality visual condition. Indeed, a paired sample *t* test showed T2 accuracy to be significantly higher at the 83-msec SOA than at the 249-msec SOA [$t(11) = 2.59$, $p < .01$]. It is not clear why visual T1 accuracy was so high at the shortest negative SOA in this crossed condition. At this SOA, T1 was presented just 83 msec before an auditory T2. As will be presented below, the visual T1 did not produce accuracy impairments for the auditory T2

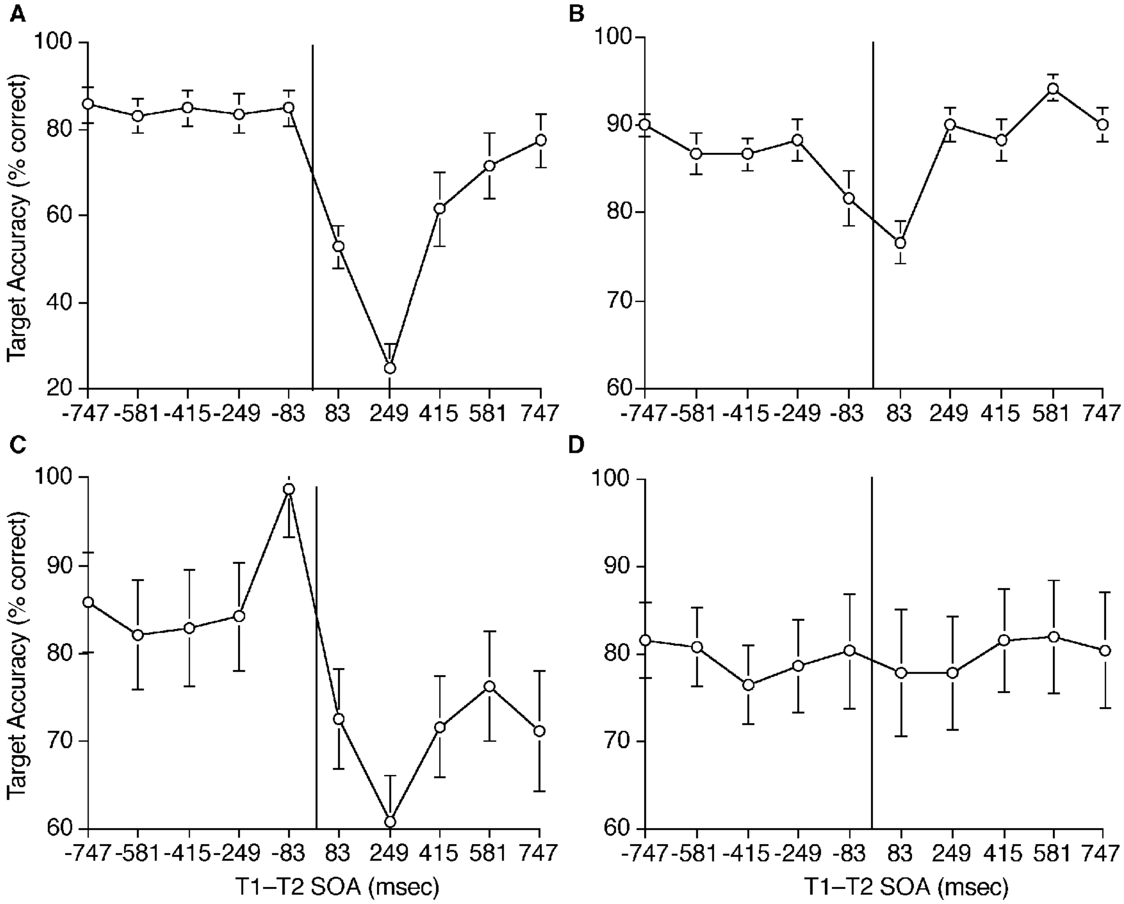

Figure 1. The Experiment 1 group mean percentage of correct target responses as a function of target number (T1 or T2) and T1–T2 stimulus onset asynchrony (SOA). Positive SOAs reflect T2 performance at each SOA after T1, and negative SOAs reflect T1 performance at each SOA backward from T2. Panel A contains means obtained in the visual within-modality condition, panel B contains means obtained in the auditory within-modality condition, panel C contains means obtained in the visual cross-modality condition, and Panel D contains means obtained in the auditory cross-modality condition. Note the smaller scale in panel A. Note also that T1 data (to the left of the line) in panel C come from the visual-T1–auditory-T2 condition and that T2 data (to the right of the line) in panel C come from the auditory-T1–visual-T2 condition. Note that T1 data (to the left of the line) in panel D come from the auditory-T1–visual-T2 condition and that T2 data (to the right of the line) in panel D come from the visual-T1–auditory-T2 condition.

at short SOAs. It is possible that, instead, the effect works in the opposite manner, so that the auditory T2 produces benefits for its cross-modal target if T2 is presented within 100 msec or so of T1.

The cross-modality auditory condition produced no significant main effect of SOA ($F < 1$), no significant main effect of target number ($F < 1$), and no SOA × target number interaction ($F < 1$). The statistics match the essentially flat line observable in Figure 1D and indicate that no AB whatsoever was observed in the cross-modality auditory condition.

Discussion

Robust ABs were observed in the within-modality visual condition, in the within-modality auditory condition, and with cross-modality visual targets. No AB was observed with cross-modality auditory targets. These results demonstrate that the AB can be found under conditions that do not encourage preparatory task-set switching. Chun and Potter (1995) demonstrated this for cases in which both targets were visual. The present results suggest that this is also true when one or both targets are auditory.

However, as was mentioned above, both T1 and T2 performances were reduced at short SOAs in the within-modality auditory condition. This makes it unclear whether the decrease in T2 accuracy at short SOAs resulted from attentional factors (true AB) or from perceptual factors. Indeed, perceptual confusion may have been particularly likely in the within-modality auditory condition, given that two tones and two targets were all presented within 100 msec of each other at the shortest SOA.

Furthermore, the within-modality auditory AB was much smaller in size than the within-modality visual AB, and the cross-modality AB was robust with visual T2s but absent with auditory T2s. This may suggest that when preparatory task-set switching is removed from designs, the AB effects are more robust for visual T2s than for auditory T2s. Indeed, there is evidence for this pattern in AB designs that do promote task-set switching (e.g., Arnell & Jolicœur, 1999; Shulman & Hsieh, 1995). However, in the present experiment, the visual identification task was essentially a 24AFC task, whereas the auditory identification task was a 2AFC task. Of course, the participants would be much less likely to guess the correct response during the visual target identification task (approximately 4%) than during the auditory identification target task (50%). Because of this, it may have been more difficult to observe reliable ABs with auditory T2s in the present experiment, and their magnitude relative to ABs with visual T2s may be underestimated. Furthermore, each auditory stimulus was played for 80 msec with only a 3.3-msec blank ISI between auditory stimuli. However, each visual stimulus was presented for only 16 msec with a 67-msec blank ISI between visual stimuli. Thus, although the item–item SOA was the same for the visual and the auditory streams, the difference in the duration/ISI ratio means that the actual presentation duration of visual targets was briefer than the actual presentation duration of auditory targets. It is possible that this also contributed to the relatively smaller ABs observed for auditory T2s. In Experiment 2, these differences between the visual and the auditory conditions were removed.

EXPERIMENT 2

The design of Experiment 2 was the same as that used in Experiment 1. However, in Experiment 2, each visual target and each auditory target was one of 4 letters, so that both auditory and visual tasks required 4AFC responses. Also, the duration/ISI ratio was more closely matched for the two streams. This experiment provided another opportunity to look for a within-modality auditory pattern that could unambiguously be interpreted as true AB, as opposed to possibly representing perceptual confusion. Also, with the number of target alternatives and presentation durations matched for visual and auditory streams, the presence/absence and relative magnitude of the ABs could be assessed with more validity.

Method

Participants. Seventeen participants (9 females) participated individually in a 1-h session. All were North Dakota State University undergraduate students participating for credit in an introductory psychology course. All the participants reported normal (or corrected-to-normal) visual acuity and normal hearing. Three participants produced less than 50% correct T1 accuracy and were, therefore, removed prior to analysis.

Stimuli and Procedure. The 2 (T1 modality) × 2 (T2 modality) × 5 (SOA) factorial design was identical to that used in Experiment 1. The stimuli were the same as those in Experiment 1, except for the following changes. Auditory targets and visual targets were now randomly drawn from a four-letter set (K, L, R, and Y). As before, T1 and T2 identity were independent, so that the identity of T1 in no way predicted the identity of T2. The target letters were no longer presented as filler letters in the visual or auditory streams. The participants were familiarized with the target sounds (K, L, R, and Y) before attempting any trials. Each visual letter was presented for 66.8 msec, followed by a 16.7-msec blank ISI, during which only the gray background was visible. As before, each auditory letter was presented to the right ear for 80 msec, followed by a 3.5-msec blank ISI, whereas the 50-msec auditory tone was presented to the left ear. Also as before, the SOA was 83.5 msec for stimuli in both the auditory and the visual streams, producing a presentation rate of almost 12.0 letters/sec. As in Experiment 1, modality combination varied randomly from trial to trial, so that the participants never knew the combination for the upcoming trial.

Results

The responses were scored as in Experiment 1. Mean target accuracy (percent correct) is plotted for T1 and T2 in Figures 2A–2D as a function of T1and T2 modality combination and T1–T2 SOA. Negative SOAs again reflect T1 performance plotted as a function of the SOA backward from T2, and positive SOAs reflect T2 performance as a function of SOA forward from T1. Chance performance equaled 25% for both the visual and the auditory tasks.

The mean target accuracy was again submitted to an ANOVA with modality (visual or auditory), target modality relationship (T1 and T2 modality crossed or within-modality), target number (T1 or T2), and T1–T2 SOA as within-subjects factors. The analysis revealed that all main effects and interactions were significant (all *p*s $< .01$), including the four-way interaction [$F(4,52) = 4.61, p < .01$].

The effects of target number (T1 or T2) and T1–T2 SOA on response accuracy were analyzed separately for within-modality and cross-modality visual targets and for within-modality and cross-modality auditory targets. The within-modality visual condition produced a significant main effect of SOA and of target number and a significant SOA × target number interaction (all *p*s $< .001$). When the effect of SOA was analyzed separately for T1 and T2, using one-way ANOVAs, a significant effect of SOA was found for T2 [$F(4,52) = 9.30, p < .001$] and for T1 [$F(4,52) = 15.08, p < .001$]. Figure 2A shows that the within-modality visual condition again produced a large and robust AB with the often characteristic U-shaped function [accuracy at the 83-msec SOA was significantly greater than accuracy at the 249-msec SOA; $t(13) = 3.15, p < .01$].

The within-modality auditory condition produced a significant main effect of SOA [$F(4,52) = 4.50, p < .01$], a significant main effect of target number [$F(4,52) = 1{,}037.19, p < .001$], and importantly, a significant SOA × target number interaction [$F(4,52) = 4.18, p < .01$]. Unlike Experiment 1, in which accuracy for both T1 and T2 was impaired at the shortest SOAs, Figure 2B (note different scale) shows that T2 accuracy significantly decreased at short SOAs [$F(4,52) = 5.61, p < .001$], with T1 accuracy remaining constant across the SOAs ($F < 1$).

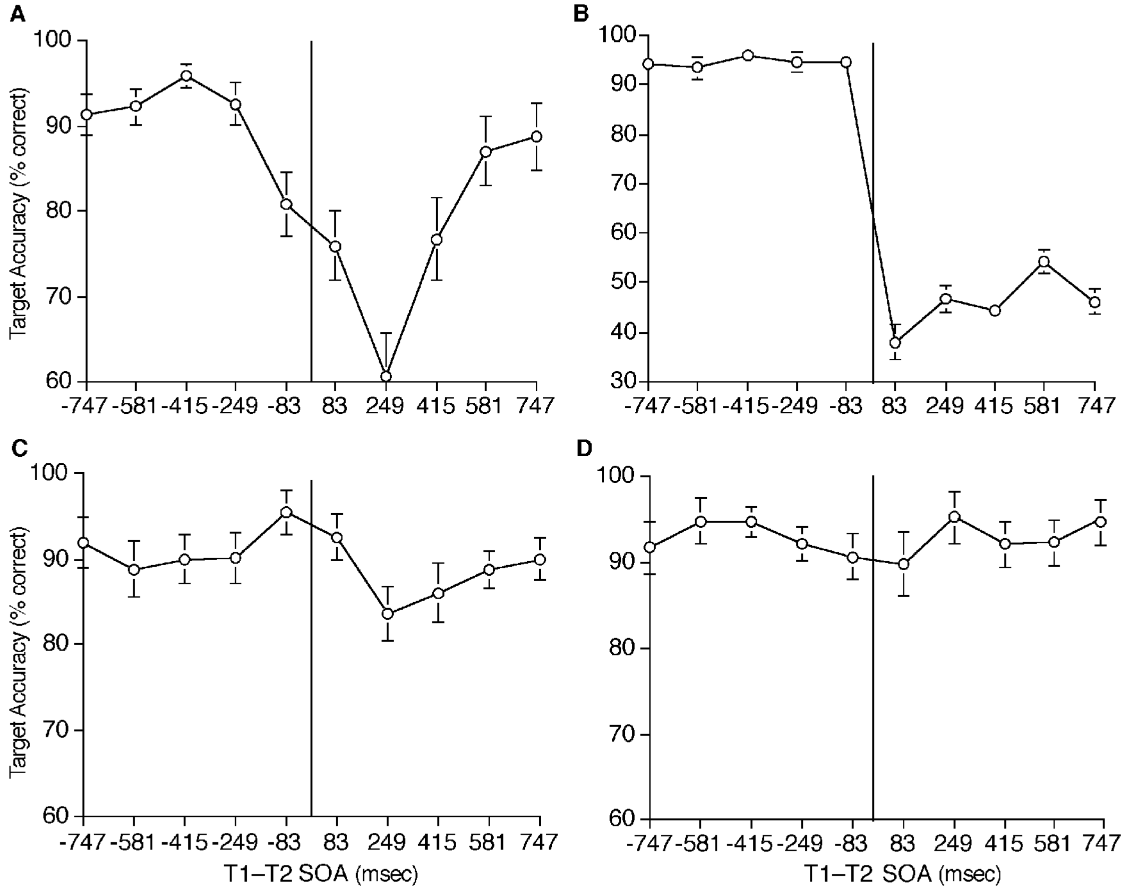

Figure 2. The group mean percentage of correct target responses from Experiment 2 as a function of target number (T1 or T2) and T1–T2 stimulus onset asynchrony (SOA). Positive SOAs reflect T2 performance at each SOA after T1, and negative SOAs reflect T1 performance at each SOA backward from T2. Panel A contains means obtained in the visual within-modality condition, panel B contains means obtained in the auditory within-modality condition, panel C contains means obtained in the visual cross-modality condition, and panel D contains means obtained in the auditory cross-modality condition. Note the smaller scale in panel B. Note also that T1 data (to the left of the line) in panel C come from the visual-T1–auditory-T2 condition and that T2 data (to the right of the line) in panel C come from the auditory-T1–visual-T2 condition. Note that T1 data (to the left of the line) in panel D come from the auditory-T1–visual-T2 condition and that T2 data (to the right of the line) in panel D come from the visual-T1–auditory-T2 condition.

In the cross-modality visual condition, there was a significant main effect of SOA [$F(4,52) = 3.98, p < .01$], but no significant effect of target number [$F(1,13) = 2.84, p > .10$] and no significant interaction of SOA and target number ($F < 1$). Analyses revealed a significant effect of SOA for T2 [$F(4,52) = 2.85, p < .05$], but no significant effect of SOA for T1 [$F(4,52) = 1.59, p > .19$]. The data from the cross-modality visual condition are presented in Figure 2C. Although quite small in size in the present experiment, the cross-modality visual AB again shows the U-shaped pattern. A paired t test indicated that T2 accuracy was significantly higher at the 83-msec SOA than at the 249-msec SOA [$t(13) = 3.30, p < .01$]. Not only was visual T2 accuracy highest at the shortest and longest SOAs in the cross-modality condition, but so was visual T1 accuracy (although not significantly, as was reported above), thus rendering the target number × SOA interaction nonsignificant. However, the lack of interaction here should not make the cross-modality visual AB suspect. Recall that the visual T2s from Figure 2C were presented on the same trials as the auditory T1s from Figure 2D, and not on the same trials as the visual T1s from Figure 2C. The cross-modality auditory T1 data show no effect of SOA (see below), and when a factorial ANOVA was performed with SOA and target number (auditory T1 and visual T2 from the same trials), the interaction was significant [$F(4,44) = 4.06, p < .05$].

The cross-modality auditory condition again produced no significant main effect of SOA ($F < 1$), no significant main effect of target number ($F < 1$), and no SOA × target number interaction ($F < 1$). The equal performance across targets and SOAs indicates that, as in Experiment 1, there was no cross-modality AB for auditory T2s. This is readily observable in Figure 2D.

Discussion

The results of Experiment 2 replicate those of Experiment 1. Reliable ABs were again observed in both within-modality conditions and in the cross-modality condition with visual T2s. As in Experiment 1, no AB was observed in the cross-modality condition with auditory T2s. Finding the same pattern of results in Experiments 1 and 2 provides strong evidence that the pattern observed in Experiment 1 did not result from differences in number of response alternatives for visual or auditory targets or from differences in presentation duration for visual and auditory stimuli. However, either reducing the number of response alternatives or lengthening the presentation duration for visual stimuli (or both) appears to have reduced the ABs for visual targets. In Experiment 1, in the within-modality visual condition, T2 performance changed almost 50% across SOAs, but in Experiment 2, T2 performance changed only about 30% across SOAs in the same condition. In Experiment 1, T2 performance changed about 15% across SOAs in the cross-modality visual condition, but it changed only about 8% in Experiment 2.

Recall that it was unclear whether the within-modality auditory accuracy pattern from Experiment 1 reflected true AB or perceptual confusion among tones and targets. In the present experiment, within-modality auditory T1 performance was stable across SOAs, providing an unambiguous interpretation of the auditory within-modality pattern as an AB. It is unclear why T1 performance varied across SOAs in Experiment 1, but not in Experiment 2, given that the only change in the auditory stimuli across experiments was the number of response alternatives. Regardless, Experiment 2 now provides a convincing demonstration of auditory AB under conditions that do not promote task-set switching.

GENERAL DISCUSSION

In both Experiments 1 and 2, ABs were observed for cross-modal dual-task presentations with visual T2s and for within-modality dual-task presentations, both when two targets were in the visual modality and when two targets were in the auditory modality. These effects were found under conditions that minimized preparatory task-set switching. T1 and T2 identity and modality were independent and varied unpredictably from target to target. Chun and Potter (1995) demonstrated a purely visual AB under conditions that did not support task-set switching. The present results suggest that an AB can also be observed without task-set switching when one or both targets are auditory.

One could speculate that although the present design does not encourage preparatory task-set switching, the participants engaged in this strategy anyway. One could also suggest that task/modality readiness was driven by stimulus occurrence, rather than by participant expectancy, in which case T1 may have shifted the state of readiness in favor of its own task. Several pieces of evidence suggest that task-set switching, whether driven by endogenous preparatory switching or exogenously by the T1 task/stimulus, does not account for the present results. First, no AB was found in either experiment for cross-modal auditory T2s, despite the fact that visual T1s preceded them. If a change in task and modality was sufficient to produce an AB-like pattern, this pattern should have been found in the cross-modality condition with auditory T2s. It was not. The presence of an AB in the within-modality conditions in which no switching was possible, and the absence of an AB in the cross-modality auditory T2 condition suggest that task switching is neither necessary nor sufficient to produce an AB pattern.

Second, the AB observed for visual T2s in the cross-modal condition showed a significant +1 sparing effect in both experiments. The term *+1 sparing* refers to a pattern in which T2 accuracy is higher when T2 comes immediately after T1 than when it comes slightly later (see Visser, Bischof, & Di Lollo, 1999, for a review of +1 sparing). Although the AB can be found without the presence of +1 sparing, +1 sparing is quite often observed in the AB paradigm and is interesting, given that most dual-tasks costs are greatest at the closest target separation. In the visual cross-modality condition of the present experiments, T2 accuracy was higher at the 83-msec SOA than at the 250-msec SOA, demonstrating +1 sparing (see Figures 1C and 2C). The presence of +1 sparing in a cross-modality condition makes it very unlikely that the T2 pattern was produced by switching, as opposed to representing true AB. Indeed, Potter et al. (1998) argued that the presence of +1 sparing could be taken as an indicator of true AB. If +1 sparing is present, Potter et al. suggested, the results represent true AB, but if +1 sparing is not present, the results represent either a combination of true AB and task switching or pure task switching. 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., even though the absence of +1 sparing tells us nothing about the validity of the AB, its presence does suggest true AB).

It is difficult to conceive of how the +1 sparing could have emerged in the visual cross-modality condition if the AB pattern was an artifact of task switching. If task-set switching (either preparatory or stimulus driven) was responsible, T2 accuracy gains should not have been seen until later SOAs, after reconfiguration had been completed. If one takes the +1 sparing effect to mean that reconfiguration has already been completed by the shortest SOA, it is not clear why accuracy was so low at the next (250-msec) SOA. If the AB pattern was due to task switching and the participants had not reconfigured their task set until after 250 msec, no +1 sparing should have been found. Also, if the AB pattern was due to task switching and the participants had reconfigured their task set by the shortest SOA, no T2 accuracy deficit should have been found. Accordingly, the presence of a

+1 sparing effect and a T2 AB-like deficit in the cross-modal condition with visual targets suggests that the AB was not an artifact of task-set switching.

Third, as was mentioned in the introduction, readiness models are characterized by complementary shifts in the efficiency of two tasks, so that as one task becomes less prepared and, therefore, is performed less efficiently, the other should become more prepared and, thus, be performed better. The present design allowed us to look for evidence of this complementarity. T1 accuracy could be taken as a baseline, given that the participants should have been equally prepared for both visual and auditory T1s at the start of each trial. Auditory target accuracy for T1 and T2 was equivalent in the cross-modality conditions of Experiments 1 and 2 (Figures 1D and 2D), in which no AB was observed. This suggests that the participants were just as prepared for an auditory T2 after viewing a visual T1 as they were prior to the presentation of T1. Furthermore, in Experiment 1 on visual-T1–auditory-T2 trials, visual T1 accuracy was significantly higher at the shortest SOA (see Figure 1C), yet the auditory T2 that immediately followed it showed no concomitant decrease in accuracy (see Figure 1D). The presence of ABs in the other modality combinations makes such interpretations slightly more difficult. However, in the within-modality visual conditions at the longest SOA (after the AB had ended), T2 performance approximated but did not exceed T1 performance in both experiments (see Figures 1A and 2A). The same can be said for the within-modality auditory condition in Experiment 1 (see Figure 1B) and for the cross-modality visual targets in Experiment 2 (Figure 2C). Indeed, only in the cross-modality condition with visual targets can one see evidence of reduced readiness to a visual T2 that followed an auditory T1 in Experiment 1 (see Figure 1C, where T1 accuracy is higher than T2 accuracy even at the longest SOA). However, even here, it is unlikely that task switching accounts for the shape of the T2 accuracy function, given that +1 sparing is observed in this condition, as was discussed above. The above patterns provide evidence against the kind of complementarity assumed in readiness models and therefore suggests that neither preparatory nor stimulus driven task-set switching can explain the present findings.

In both experiments, ABs were observed in both within-modality conditions and the cross-modality condition with visual T2s. Contrary to the arguments of Potter et al. (1998), these results suggest that the AB is not uniquely visual in nature and that the AB can be found when one or both targets are not visual, using a design in which task-set switching was not likely. However, the magnitude of the AB did not appear to be equal in the four conditions. Even after the number of response alternatives and presentation durations had been equated for visual and auditory stimuli (Experiment 2), the AB appeared largest in the within-modality visual condition and smaller in the within-modality auditory and cross-modality visual T2 conditions and was absent in the cross-modality auditory T2 condition. Arnell and Duncan (2002), Potter et al. (1998, Experiment 4), and Shulman and Hsieh (1995) have all shown apparently larger ABs within modality/task than across modality/task. Also, previous work on the cross-modality AB (e.g., Arnell & Jolicœur, 1999; Shulman & Hsieh, 1995) has shown apparently larger ABs for visual T2s than for auditory T2s, possibly owing to the greater temporal resolution of the auditory system or the relatively greater capacity and/or longer duration of echoic memory. So, the AB may be larger when targets are presented within the same modality, as compared with across modality, and when T2 is visual, as compared with auditory. Combining these factors leads to a pattern in which the within-modality visual condition can be expected to have the largest AB and the cross-modality auditory T2 condition can be expected to have the smallest AB. This pattern is also observed here: The AB in the visual within-modality condition appears to be the largest, the ABs in the auditory within-modality condition and the visual cross-modality condition appear smaller, and the auditory cross-modality condition shows no AB. Therefore, it seems that removing preparatory task-set switching has not eliminated the auditory and cross-modal AB and has preserved the relative size of the ABs found in many of the previous cross-modality AB studies that encouraged task-set switching.

Finding a cross-modal AB that is not the result of task-set switching suggests that the processing limitations that underlie the AB are central (amodal) in nature, at least in part. As was suggested by Arnell and Duncan (2002), larger ABs within tasks and/or modalities, as compared with across tasks and/or modalities, can be explained by a model in which within-task/modality AB results from both within and central sources of interference and cross-task/modality AB results only from central sources of interference. It is possible that the central resource requirements were particularly high in the present experiments, given the fast presentation rate and the need to monitor concurrently both the auditory and the visual streams. This high central resource demand may be in contrast to that required by the participants in Potter et al. (1998), in which a slower presentation rate was used and participants were required to monitor only one stream at a time.

To the extent that some of the within-modality interference is perceptual in nature, not attentional, it seems prudent to use cross-modality targets (preferably, visual T2s) in future experiments. Also, to the extent that task-set switching confounds the results of AB experiments, it seems prudent to use designs, like the present one, that discourage and measure task-set switching in future experiments.

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). Separate and shared sources of dual-task cost in stimulus identification and response selection. *Cognitive Psychology*, **44**, 105-147.

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

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., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavioral Research Methods, Instruments, & Computers*, **25**, 257-271.

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.

Jolicœur, P. (1999). Restricted attentional capacity between sensory modalities. *Psychonomic Bulletin & Review*, **6**, 87-92.

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

Miller, M. 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.

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 operation. *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.

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 attention blink in mixed modality streams*. Paper presented at the 36th Annual Meeting of the Psychonomic Society, Los Angeles.

Soto-Faraco, S., Spence, C., Fairbank, K., Kingstone, A., Hillstrom, A. P., & Shapiro, K. L. (in press). A crossmodal attentional blink between vision and touch. *Psychonomic Bulletin & Review.*

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

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

(Manuscript received June 20, 2000;
revision accepted for publication September 12, 2001.)