

VU Research Portal

The absence of an auditory-visual attentional blink is not due to echoic memory.

van der Burg, E.; Olivers, C.N.L.; Bronkhorst, A.W.; Koelewijn, T.; Theeuwes, J.

published in

Perception & Psychophysics

2007

DOI (link to publisher)

[10.3758/BF03193958](https://doi.org/10.3758/BF03193958)

document version

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

citation for published version (APA)

van der Burg, E., Olivers, C. N. L., Bronkhorst, A. W., Koelewijn, T., & Theeuwes, J. (2007). The absence of an auditory-visual attentional blink is not due to echoic memory. *Perception & Psychophysics*, 69(7), 1230-1241. <https://doi.org/10.3758/BF03193958>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

The absence of an auditory–visual attentional blink is not due to echoic memory

ERIK VAN DER BURG AND CHRISTIAN N. L. OLIVERS
Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

ADELBERT W. BRONKHORST
Vrije Universiteit Amsterdam, Amsterdam, The Netherlands
and TNO Human Factors, Soesterberg, The Netherlands

AND

THOMAS KOELEWIJN AND JAN THEEUWES
Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

The second of two targets is often missed when presented shortly after the first target—a phenomenon referred to as the attentional blink (AB). Whereas the AB is a robust phenomenon within sensory modalities, the evidence for cross-modal ABs is rather mixed. Here, we test the possibility that the absence of an auditory–visual AB for visual letter recognition when streams of tones are used is due to the efficient use of echoic memory, allowing for the postponement of auditory processing. However, forcing participants to immediately process the auditory target, either by presenting interfering sounds during retrieval or by making the first target directly relevant for a speeded response to the second target, did not result in a return of a cross-modal AB. The findings argue against echoic memory as an explanation for efficient cross-modal processing. Instead, we hypothesized that a cross-modal AB may be observed when the different modalities use common representations, such as semantic representations. In support of this, a deficit for visual letter recognition returned when the auditory task required a distinction between spoken digits and letters.

To investigate the temporal dynamics and limitations of the human visual system, a rapid serial visual presentation (RSVP) task is often used. Within RSVP tasks, multiple stimuli are rapidly presented one at a time. When participants are asked to identify and report two target stimuli (T1 and T2) from this rapid stream, they often show a deficit in identifying T2 when T2 is presented within half a second after T1. This temporal limitation of the visual system is known as the *attentional blink* (AB; Raymond, Shapiro, & Arnell, 1992).

Shapiro, Raymond, and Arnell (1994) proposed that the AB phenomenon is a purely visual limitation. However, recent studies have generated compelling evidence for the same limitation when two target stimuli are auditory targets within a rapid serial auditory presentation (RSAP), indicating that an AB can be found in modalities other than the visual one (Arnell & Jenkins, 2004; Arnell & Jolicœur, 1999; Duncan, Martens, & Ward, 1997; Hein, Parr, & Duncan, 2006; Potter, Chun, Banks, & Muckenhaupt, 1998; Shen & Mondor, 2006; Soto-Faraco & Spence, 2002; Tremblay, Vachon, & Jones, 2005). Furthermore, Hillstrom, Shapiro, and Spence (2002) observed an AB within the tactile modality, further supporting the idea that the AB phenomenon is not a pure visual limitation.

An important question is, then, whether an AB can also occur between modalities. Duncan et al. (1997) and, more recently, Hein et al. (2006) and Soto-Faraco and Spence (2002) failed to find an AB when T1 and T2 were from different modalities (auditory–visual or visual–auditory), despite clear ABs when the targets were from the same modality. Duncan et al. therefore proposed that the AB deficit is a modality-specific restriction.

In contrast, Arnell and Jolicœur (1999), as well as Potter et al. (1998), were more successful in observing an AB deficit when T1 and T2 were from different modalities (auditory–visual and visual–auditory). In Arnell and Jolicœur's study, participants were asked to identify a digit (T1) and to detect the presence of a letter "X" (T2) among visual and spoken distractor letters. The reported cross-modal effects suggest that the AB involves a central amodal processing limitation. Furthermore, Soto-Faraco et al. (2002) observed an AB between the visual and tactile modalities, further supporting the idea that the AB reflects an amodal bottleneck.

Potter et al. (1998) replicated the cross-modal AB when participants were asked to report a digit in one modality (auditory or visual) followed by a letter in the other modality (following Arnell and Jolicœur's task). However, T2

remained unaffected when both targets were digits among distractor letters. Potter et al. concluded that the cross-modal AB was absent when the two targets belonged to the same task set and that the cross-modal AB is therefore due to task set reconfiguration.

In response to Potter et al.'s (1998) conclusion that a task set switch is important to perceive a cross-modal AB, Arnell and Jenkins (2004) recently argued that rather than task set reconfiguration, the match between the alphanumeric class of T2 and the distractors plays a crucial role in the occurrence of an AB. Arnell and Jenkins found that when the targets were letters among spoken and visual distractor letters, an AB was present, while a task-set switch was absent. In contrast, when the targets were digits among spoken and visual distractor letters, an AB was absent.

An Alternative Hypothesis: Echoic Memory

All in all then, the evidence regarding cross-modal ABs is rather mixed and inconsistent and has led to various accounts. So far, however, an alternative explanation for the variety of results has been unexplored—namely, one in terms of echoic memory. Evidence reviewed by Cowan (1984) suggests that echoic memory is an auditory memory that can be divided into two types. Information in the *short auditory store* decays within about 200–300 msec and serves close-to-immediate recognition processes. Auditory information in the *long auditory store*, on the other hand, is preserved for at least several seconds (see also Crowder, 1993; Kallman & Massaro, 1979) and allows for the retrieval of an entire auditory sequence. Importantly, such longer term storage of sequences of sounds is aided by manipulations that create stream-like percepts (see, e.g., Bregman, 1978; Dorman, Cutting, & Raphael, 1975). On the other hand, storage and/or retrieval is hindered by the presentation of similar masking sounds after the sequence stage (Crowder, 1993; Deutsch, 1970).

The existence of echoic memory, and especially the long-term version of it, raises the possibility that in cross-modal AB tasks, participants store crucial parts of the auditory stream in the auditory buffer and postpone processing of the auditory target until after they have completed the visual task. They then reconstruct the auditory target from the echoic buffer, resulting in high accuracy not only for T1, but also for T2 (see also Arnell, 2006, as well as Chun & Potter, 2001, for this suggestion). Such a strategy may be particularly effective when relatively simple auditory stimuli are used with stream-like properties, so that they can be easily stored as a whole in echoic memory. In contrast, when more complex or meaningful auditory stimuli (e.g., letters or digits) are used, observers may be forced to treat them as separate entities and may process the auditory target immediately, and thus T2 will be affected when T2 is presented shortly after T1. Indeed, consistent with this suggestion, clear cross-modal ABs were observed in studies that used meaningful auditory stimuli, such as letters and digits (see, e.g., Arnell & Jenkins, 2004; Arnell & Jolicœur, 1999; Arnell & Larson, 2002; Potter et al., 1998), whereas no cross-modal ABs

(see, e.g., Duncan et al., 1997; Hein et al., 2006) or largely reduced cross-modal ABs (Arnell & Jolicœur, 1999, Experiment 3) were observed in studies that used relatively simple auditory stimuli, such as pure tones—stimuli that may be easily processed as a single stream rather than as individual objects.

Some studies have reported an AB while using simple auditory stimuli, such as beeps (Hein et al., 2006; Shen & Mondor, 2006; Vachon & Tremblay, 2005) or short words among nonsense syllables, for which detection may have been more of an acoustic than of a semantic nature (Duncan et al., 1997; Tremblay et al., 2005). This suggests that echoic memory may not be that useful even with simple stimuli. However, note that in these studies, both the T1 and the T2 tasks involved auditory streams. Thus, the interference may be of a more direct, auditory nature, rather than involving the supposed central limitations one is interested in when studying cross-modal ABs. Moreover, it would make little sense to try and perform the task on an echoic trace of the stimulus when both T1 and T2 are auditory in nature, since the fidelity of the trace would presumably be much lower than that of the real stimulus. Echoic memory is more useful in the cross-modal case, where the percept from one modality can be temporarily stored while the other modality is processed. This is the case we are interested in here.

The aim of the present article is to investigate whether the contrasting results with regard to the appearance of an auditory–visual AB are indeed due to auditory target processing being postponed until after the termination of the visual task. In principle, our echoic memory hypothesis also holds for the reverse case, in which an auditory target follows the visual target. However, the visual–auditory AB is typically weaker than the auditory–visual AB (which in itself is not too strong either; see, e.g., Arnell & Jolicœur, 1999; Arnell & Larson, 2002; Potter et al., 1998), leaving relatively little room for experimental manipulations. Hence, we focused on the auditory–visual AB only.

In Experiment 1, we established an AB within the auditory and visual domains, to make sure that the auditory target (T1) we used can induce a blink and that the visual target (T2) we used can be blinked. In Experiment 2, we used the (auditory) T1 and the (visual) T2 from Experiment 1 to try to generate an auditory–visual AB. We failed. In Experiments 3 and 4, we investigated whether this failure was due to postponed auditory T1 processing, by forcing participants to process the auditory target immediately, so that efficient use of echoic memory was discouraged or even impossible. In Experiment 3, interfering distractor tones were presented during the response stage, which would presumably interfere with echoic retrieval. Experiment 4 made T1 directly relevant for the T2 task, and thus, by necessity, participants needed to process T1 before T2. Still, none of these manipulations generated an AB, making an echoic memory explanation an unlikely candidate. In Experiments 5 and 6, we tested the alternative hypothesis that in order to find interference, the auditory and visual stimuli need to be from the same class of stimuli (cf. Arnell & Jenkins, 2004). In support of this, an

AB was observed when both the auditory and the visual stimuli were alphanumeric in nature.

EXPERIMENT 1 Establishing Within-Modality Blinks

Before we can test for an auditory–visual AB, we first need to establish the occurrence of an AB within each of the modalities, to make sure that the visual T2 can, in principle, be blinked and that the auditory T1 can, in principle, induce such a blink. Experiment 1A tested for a visual AB by presenting participants with an RSVP stream and requiring them to detect either a letter “b” or a “d” as the first target and to detect the presence of an “X” as the second target. In the control condition, only the “X” needed to be detected. Experiment 1B tested for an auditory AB by presenting participants with an RSAP task involving a judgment of pitch for the first target tone and a judgment of direction for the second target tone. In the control condition, only the second target needed to be judged. Within-modality ABs would be expressed as worse performance for the second target after processing the first target within each of the modalities, relative to the respective control performance.

Method

Participants

Twelve students (10 male; mean age 20.8 years, ranging from 17 to 35 years) participated in Experiment 1A and 10 new students (4 male; mean age 22.0 years, ranging from 18 to 32 years) participated in Experiment 1B as paid volunteers. Each participant received €3.5 for a single 30-min session.

Apparatus and Stimuli

The experiment was run in a dimly lit, air-conditioned cubicle. Participants were seated at a distance of approximately 80 cm from the monitor. The visual stimuli included all the letters of the alphabet except the letters W, N, F, and S (see also Arnell & Jolicœur, 1999; Potter et al., 1998). All letters (except the letters “b” and “d”) were capitals and were presented in a black 48-point Geneva font (0.63 cd/m², 1.4° width and 1.6° height) at the center of a gray background (9.34 cd/m²). The first target (T1) was either the letter “b” or “d,” one of which was always present. The second target (T2) was the letter “X,” which was present on 50% of the trials.

In the auditory experiment, participants wore headphones (Sennheiser HD202). The auditory distractors were randomly selected from a set of eight different tones, ranging from 400 to 800 Hz, equally spaced on a logarithmic scale. The first target (T1) was one of two tones, both higher in frequency than the distractor tones. The highest of the two possible target tones was drawn from a set of six different tones that were equally spaced on a logarithmic scale, ranging from 1767 to 3200 Hz, starting at a frequency of 2378 Hz. The lowest of the two tones was kept constant at 1600 Hz. To keep T1 performance at a reasonable level, the frequency increased by one step when T1 accuracy was lower than 78% and decreased by one step when T1 accuracy was higher than 91%. This was done after each block of 32 trials. It turned out that during the experiments reported here, T1 performance was close to optimal (i.e., close to or even exceeding 91%) even for the smallest difference between target tones. The distractor tones and T1 were presented to both ears. The second target (T2) was a randomly selected distractor tone that was perceptibly presented to the left or to the right by reducing the amplitude in the contralateral ear by 14 dB. All individual letters or tones were presented for 16 msec, followed by an 80-msec blank interval,

which resulted in a stimulus onset asynchrony (SOA) of 96 msec between the different elements in the streams (10.4 elements/sec).

Design and Procedure

Each trial began with a fixation cross presented for 500 msec at the center of the screen followed by a blank screen for the same period. Stimulus streams (whether auditory or visual) started with a randomly determined number of distractors (4, 6, 8, 10, or 12), before T1 was presented. T2 then followed at four possible lags (lag 1, 2, 3, or 8). Note that in both visual and auditory streams, the first target was followed by a total of nine elements, such that the second target was always followed by at least one distractor. Participants were asked to report both T1 and T2 in the dual-task condition and only T2 in the control condition. In Experiment 1A, after the termination of the sequence, participants were asked to make an unsped response to T1 by pressing the b or d key when they identified the letter “b” or “d,” respectively, and to make an unsped response to T2 by pressing the 0 or 1 key when the letter “X” was absent or present, respectively. In Experiment 1B, after the termination of the sequence, participants were asked to make an unsped response to T1 by pressing the 5 or 6 key when they identified the high or very high target tone, respectively, and to make an unsped response to T2 by pressing the l or r key when the second target tone was presented left or right, respectively. Participants initiated the next trial by pressing the space bar. Participants received instructions prior to the experiment, which emphasized accuracy. In Experiment 1B, participants first performed three blocks of 20 random trials in which they only practiced the T2 task. Participants were allowed to participate in Experiment 1B only when T2 control performance exceeded 80%. In both experiments, there were two practice blocks (one control, one dual task) of 16 random trials each. After the practice blocks, participants completed five dual-task blocks and five control blocks of 32 trials each. Blocks were presented in an alternating and counterbalanced order. Lag was randomly mixed within blocks. Participants received feedback about their overall mean accuracy after each block.

Results and Discussion

Experiment 1A

Data from practice blocks were excluded from further analysis. The data for T1 accuracy were subjected to a repeated measures univariate ANOVA, with lag (1, 2, 3, and 8) as a within-subjects factor, and alpha set at .05. The same was done with T2 accuracy with lag and task (control and dual task) as within-subjects factors for those trials on which T1 was correctly identified and T2 was present. The reported values for MS_e and p are those after a Huynh–Feldt correction for sphericity violations. In subsequent experiments, unless otherwise stated, all results were analyzed in the same way.

T1 accuracy. Overall, T1 accuracy was at 88.1% correct. The ANOVA yielded no significant effect of lag ($F < 1$).

T2 accuracy. Overall, T2 false alarm rate (on T2-absent trials) was at 12.1% in the control condition and at 11.7% in the dual-task condition. Figure 1 presents the mean percentage correct for T2 on T2-present trials as a function of lag and task. In the dual-task condition, mean percentage correct for T2 was calculated for those trials on which participants identified T1 correctly.

There was a significant main effect of task [$F(1,11) = 102.20$, $MS_e = .031$, $p < .001$]. T2 accuracy was better when participants were asked to detect only the “X”

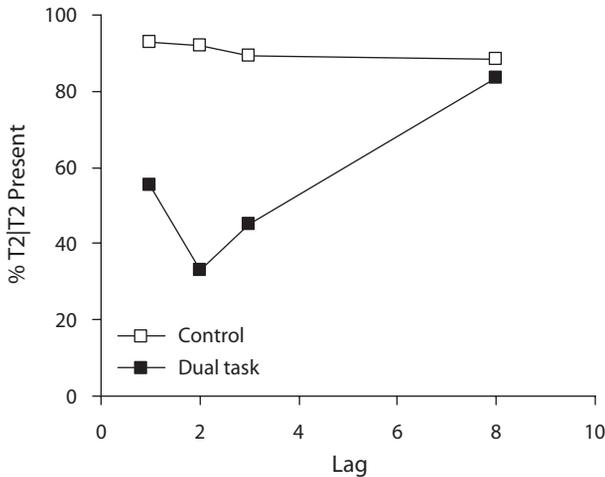


Figure 1. Experiment 1A: Mean percentages correct for a visual T2, given that an auditory T1 was correctly identified (in the dual-task condition) and T2 was present, as a function of lag and task. Note that here (as in all other figures), the bottom of the scale was adjusted to the chance level of correct response.

(90.6%) than when participants were asked to identify T1 and to detect the “X” (54.2%). There was also a significant main effect of lag [$F(3,33) = 21.99$, $MS_e = .014$, $p < .001$]. Overall, performance increased with increases in lag. Moreover, there was a significant two-way interaction between lag and task [$F(3,33) = 28.89$, $MS_e = .011$, $p < .001$]. The interaction was further analyzed in detail by pairwise t tests for each lag (1–3 and 8). These revealed significant differences at lags 1–3 (all p values $< .001$), but not at lag 8 ($p = .224$). The U-shaped performance curve represents a classic AB pattern within the visual modality, corroborating earlier work (e.g., Arnell & Jolicœur, 1999; Duncan et al., 1997; Hein et al., 2006; Raymond et al., 1992).

Experiment 1B

In the present experiment, the data were analyzed in the same way as in Experiment 1A, except that there were no false alarms because T2 was always present.

T1 accuracy. Overall, T1 accuracy was at 88.3% correct. The ANOVA yielded no significant effect of lag ($F < 1.2$).

T2 accuracy. Figure 2 presents the mean percentage correct for T2, given that T1 was correctly identified (if applicable), as a function of lag and task.

There was a significant main effect of task [$F(1,9) = 12.62$, $MS_e = .014$, $p < .01$]. T2 accuracy was better when participants were asked to identify only T2 (90.2%) than when participants were asked to identify both T1 and T2 (80.7%). Performance increased slightly with increases in lag [$F(3,27) = 4.04$, $MS_e = .006$, $p < .05$]. Moreover, there was a significant two-way interaction between lag and task [$F(3,27) = 3.67$, $MS_e = .004$, $p < .05$]. The interaction was further analyzed in detail by pairwise t tests for each lag (1–3 and 8). The t tests revealed significant differences at lags 1–3 [$t(9) = 3.34$, $p < .01$; $t(9) = 3.04$, $p < .05$; $t(9) = 2.93$, $p < .05$], but not at lag 8 ($p = .68$). This pattern of results

clearly represents an AB within the auditory modality, corroborating earlier work (e.g., Arnell & Jolicœur, 1999; Duncan et al., 1997; Hein et al., 2006; Shen & Mondor, 2006; Tremblay et al., 2005; Vachon & Tremblay, 2005).

EXPERIMENT 2 Absence of an Auditory–Visual AB

The findings of Experiment 1 are important in that they show that the visual target (T2) used in Experiment 1A can, in principle, be affected by an earlier target and that the auditory T1 used in Experiment 1B can, in principle, induce an AB. In the present experiment, we therefore used the visual target (T2) from Experiment 1A and the auditory target (T1) from Experiment 1B in an auditory–visual AB paradigm to see whether we could generate a cross-modal AB. In the dual-task condition, participants were asked to first identify which of the two high tones was presented within the auditory stream and then to detect the presence of a letter “X” in the visual stream. In the control condition, participants were asked to detect only the presence of a letter “X” in the visual stream (following Arnell & Jolicœur, 1999, Experiment 3). If processing the auditory target induces a cross-modal AB, performance for the visual target should be affected as a function of lag.

Method

The present experiment was identical to Experiment 1, except for the following modifications.

Participants

Twelve new students (6 male; mean age 19.6 years; ranging in age from 16 to 24 years) participated in the experiment as paid volunteers.

Design and Procedure

Participants received concurrently an RSAP and an RSVP stream. In the dual-task condition, participants were asked to identify the target tone (T1; high or very high in frequency) in the RSAP stream and to determine the presence of the target letter “X” (T2) in the RSVP

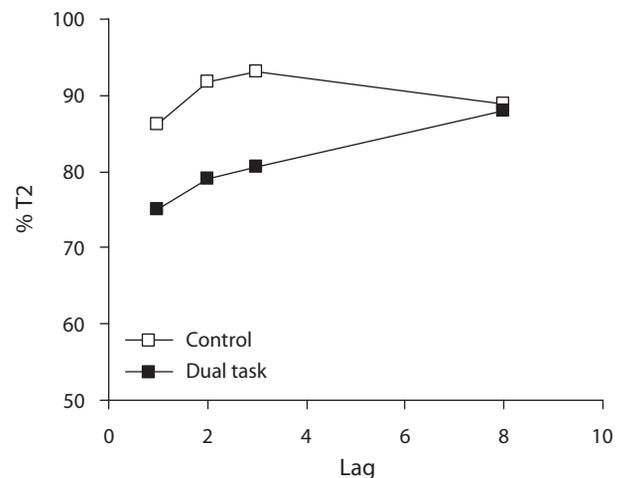


Figure 2. Experiment 1B: Mean percentages correct for an auditory T2, given that an auditory T1 was correctly identified (in the dual-task condition), as a function of lag and task.

stream. T2 appeared at eight possible lags (lags 1–8). In the control condition, participants were asked only to determine the presence of the target letter “X” (T2) in the visual stream and to ignore the auditory stream.

There were two practice blocks of 16 trials each. After the practice blocks, participants completed five dual-task blocks and five control blocks of 64 trials each. Blocks were presented in alternating and counterbalanced order.

Results and Discussion

T1 Accuracy

Overall, T1 accuracy was at 92.2% correct. The ANOVA yielded no significant effect of lag ($F < 1$).

T2 Accuracy

Overall, T2 false alarm rate (on T2-absent trials) was at 7.6% in the control condition and at 9.7% in the dual-task condition. Figure 3 presents the mean percentage of correct detection of T2 as a function of lag and task on T2-present trials for correctly identified T1s (when applicable).

There was no significant main effect of task ($F < 1$). Furthermore, there was a significant main effect of lag [$F(7,77) = 2.72$, $MS_e = .007$, $p < .05$], since T2 performance decreased slightly with increases in lag. The two-way interaction between lag and task failed to reach significance ($F = 1.4$).

The results show that extracting a tone from an auditory stream does not affect detection of a target letter in a visual stream, even though we used the same auditory and visual tasks as in Experiment 1. Participants performed just as accurately in the dual-task condition as in the control condition. The absence of an auditory–visual AB in the present experiment is consistent with the results of earlier studies using relatively simple auditory stimuli (Duncan et al., 1997; Hein et al., 2006), whereas others, using more complex semantic stimuli, have shown a cross-modal AB (Arnell & Jenkins, 2004; Arnell & Jolicœur, 1999; Potter et al., 1998).

One possible explanation is that the relatively simple auditory stimuli used here and in other studies failing to find a cross-modal AB allowed observers to first focus on the visual task and then replay the auditory stream from echoic memory, because of the long persistence (Cowan, 1984; Crowder, 1993). Such postponement of auditory processing may be more difficult with more complex or semantic auditory stimuli as used by those that did find a cross-modal deficit. The next two experiments were therefore designed to prevent participants from relying on their echoic memory and to see whether a cross-modal AB could be obtained.

EXPERIMENT 3 Interfering With Echoic Memory

The aim of Experiment 3 was to force participants to engage in immediate auditory target processing and to see whether this would generate a cross-modal AB. For this purpose, the auditory stream of distractors continued beyond the visual stream and kept on running during the entire response period. It is known that the presence of

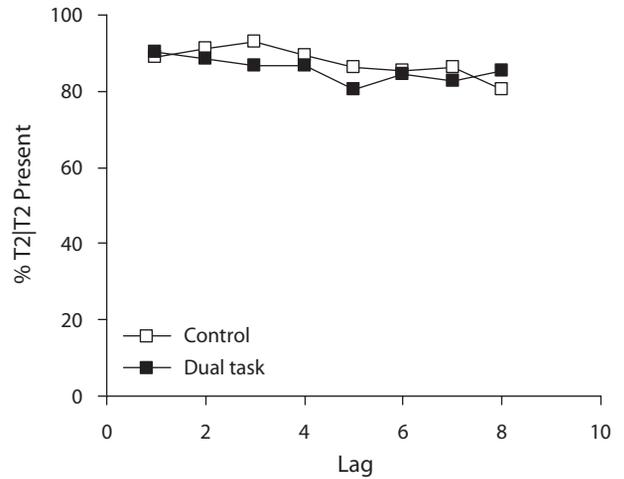


Figure 3. Experiment 2: Mean percentages correct for a visual T2, given that an auditory T1 was correctly identified (when applicable) and T2 was present, as a function of lag and task.

auditory stimuli interferes with the content of, or the retrieval from, echoic memory, especially if the two types of information are very similar (Crowder, 1993; Deutsch, 1970). As an example from everyday life, trying to retrieve a specific tune is difficult with the radio still on. We assumed that the continuous presence of distractor beeps during the response stage should discourage participants from using their auditory memory and should urge them to immediately process the auditory target when presented. If echoic memory indeed contributes to the absence of a cross-modal AB, then we might see this AB return here.

Method

Twelve new students (6 male; mean age, 19.3 years; ranging from 17 to 22 years) participated in the experiment as paid volunteers. The present experiment was identical to Experiment 2, except that participants were asked to make a response within 5,000 msec after the presentation of the last element in the RSVP stream. During this 5,000-msec period, the RSVP stream continued with random distractor tones (SOA = 96 msec).

Results and Discussion

Trials on which no response at all was made were excluded from further analysis (0.26%).

T1 Accuracy

Overall, T1 performance was at 95.4% correct. The ANOVA yielded no significant effect of lag ($F < 1$).

T2 Accuracy

Overall, T2 false alarm rate was at 8.2% in the control condition and at 5.1% in the dual-task condition. Figure 4 presents the mean percentage of correct detection of T2, for those trials on which the target tone (T1) was correctly identified (if applicable) and T2 was present, as a function of lag and task.

There was a trend toward a main effect of task [$F(1,11) = 4.58$, $MS_e = .033$, $p = .056$], reflecting the somewhat

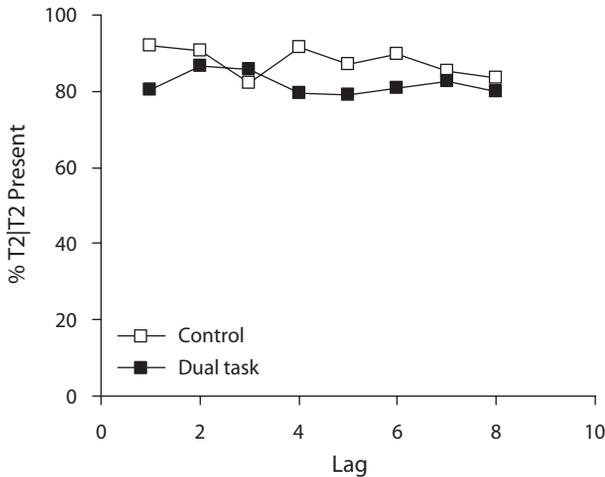


Figure 4. Experiment 3: Mean percentages correct for a visual T2, given that an auditory T1 was correctly identified (when applicable) and T2 was present, as a function of lag and task.

worse performance overall in the dual task (82.7%), as compared with the control task (88.3%). The main effect of lag failed to reach significance [$F(7,77) = 1.19$, $MS_e = .008$, $p = .32$]. The two-way interaction between lag and task was significant [$F(7,77) = 2.20$, $MS_e = .005$, $p < .05$]. t tests yielded a significant effect between the dual task and the control task for lags 1, 4, and 6 (all p values $< .05$). The differences on other lags failed to reach significance (all $ps > .1$).

The presence of a significant interaction between lag and task would suggest the presence of a cross-modal AB, providing evidence for the echoic memory hypothesis. However, a look at Figure 4 makes this hypothesis less than convincing. Although the relative deficit at lag 1 may be consistent with an AB pattern, the equally large deficits at lags 4 and 6 are not. Moreover, the interaction appears equally aided by an inexplicable dip in performance at lag 3 in the control condition. In fact, when only the dual-task condition was tested, there was no effect of lag [$F(7,77) = 1.30$, $MS_e = .008$, $p = .27$]. In any case, the pattern here does not resemble the clear cross-modal deficits found by others (Arnell & Jenkins, 2004; Arnell & Jolicœur, 1999; Potter et al., 1998) nor the unimodal blinks found in Experiment 1.

On the assumption that the presence of distractor tones during response would have interfered within echoic memory, we then conclude, on the basis of the absence of such interference effects, that observers did not use echoic memory and that it therefore does not provide a useful explanation for the absence of a cross-modal AB. Of course, one could argue that our assumption that our manipulation would prevent observers from using echoic memory was wrong in the first place. Perhaps the distractor tones indeed interfered with echoic memory, but this did not prevent observers from using it, perhaps for lack of a better strategy. However, if this were the case, we would expect a deterioration in T1 performance, relative to the previous experiments. There was no such deterioration. If

anything, T1 performance was slightly better in Experiment 3 (95.4%) than in Experiment 2 (92.2%) and Experiment 1B (88.3%).

Another possibility is that our manipulation failed altogether in that it did not interfere with echoic memory at all, allowing observers to make full use of it. Although we regard this possibility as unlikely, given the literature on echoic memory (Crowder, 1993; Deutsch, 1970), we cannot completely exclude it. We therefore designed Experiment 4 to further prevent participants from using echoic memory and to provide converging evidence as to whether echoic memory should or should not be regarded as a key contributor to efficient cross-modal processing.

EXPERIMENT 4 Using a Go–No–Go Task to Force Immediate Processing

Experiment 4 introduced another method to prevent participants from postponing auditory processing until after the visual target. The crucial condition was the go–no–go condition. In this condition, the nature of T1 determined whether T2 should be responded to or not. Furthermore, responses to T2 (when indicated by T1) were speeded—that is, participants responded as quickly as possible. Responses to T1 itself were not speeded: Participants reported T1 at the end of the trial, after T2. Note that, this way, T1 required immediate processing without its requiring an immediate response. This method presents an advantage over earlier studies that also forced immediate T1 processing (Arnell & Duncan, 2002; Dell’Acqua, Jolicœur, Pesciarelli, Job, & Palomba, 2003). Interestingly, in these studies, a cross-modal deficit for T2 was observed. However, note that by asking observers to immediately respond to T1, a strong response selection component is introduced, and the cross-modal deficit observed may therefore reflect a response- or motor-related bottleneck (Pashler, 1989, 1994). The present procedure circumvents this potential confound.

Performance in the go–no–go condition was compared with that in a dual-task condition and a control condition comparable to those in Experiment 2. If the absence of a cross-modal AB in the dual-task condition is due to the postponement of T1 processing, we should see such an AB reemerge in the go–no–go condition (in which postponement is not possible), relative to the standard dual-task condition (in which postponement could be possible) and the control condition.

Method

The present experiment was the same as Experiment 2, except for the following changes.

Participants

Fourteen new students (6 male; mean age 21.5 years; ranging from 18 to 30 years) participated in the experiment as paid volunteers. Each participant received €12 for the 100-min single session. Data from 2 participants were excluded from further analyses because they responded on more than 50% of the trials before the onset of T2.

Design and Procedure

There was an additional go–no-go task. Within this condition, participants were asked to make a speeded response as soon as T2 appeared, but only if T1 was a “go” tone (e.g., a high target tone). In the case of a “no-go” tone (e.g., a very high target tone), participants were asked to withhold their response even when T2 appeared. On 80% of the trials T1 was a go tone, and on 20% of the trials T1 was a no-go tone. The assignment of the high or the very high tone to the go or the no-go condition was counterbalanced across participants. Participants were asked to report T1 unspeeded after the presentation of the two streams in all conditions. In all conditions (control task, dual task, and go–no-go task), T2 was present on 75% of the trials and absent on 25% of the trials. Participants were also instructed to make a speeded response by pressing the space bar when they detected T2 (the visual letter “X”; in the go–no-go condition, this should be done only on go trials).

Prior to the experimental blocks, there were four practice blocks of 16 trials each. One block was used to practice the difference between the two target tones in the auditory stream. The other three practice blocks were used to practice each of the three main tasks. After the practice blocks, participants completed five control blocks, five dual-task blocks, and five go–no-go task blocks of 64 trials each, in counterbalanced order. Participants received feedback after each block on their overall mean accuracy, the number of trials on which they had responded before the presentation of the target letter, and their overall mean reaction time (RT).

Results and Discussion

Trials on which participants responded to T2 but did so too late ($RT > 1,000$ msec) were excluded from further analysis (1.5%). Trials on which participants responded too early ($RT < 300$ msec) were also excluded (8.2%).

T1 Accuracy

The data of T1 accuracy were subjected to a repeated measures ANOVA, with lag and task (dual task, go–no-go task) as within-subjects factors. Overall, T1 accuracy was at 92.4% correct for the dual task and at 93.9% correct for the go–no-go task [$F(1,11) = 3.44$, $MS_e = .003$, $p = .091$]. The analysis yielded no significant main effect of lag [$F(7,77) = 2.17$, $MS_e = .004$, $p = .112$]. The two-way interaction between task and lag also failed to reach significance ($F < 1$). Also, within the go–no-go task, there was no difference in performance on go (93.2%) and no-go (96.4%) trials, when $t < 1$.

T2 Accuracy

Overall, T2 false alarm rate was at 10.2% in the control task, at 15.8% in the dual task, and at 12.0% in the go–no-go task. Furthermore, when the target tone was a no-go tone in the go–no-go task, participants made more errors when T2 was present (13.4%) than when T2 was absent (2.6%). Figure 5 presents the mean percentage of correct detection of T2, as a function of task and lag for those trials on which T1 was a go tone, T1 was correctly identified (when applicable), and T2 was present.

There was a significant main effect of task [$F(2,22) = 3.76$, $MS_e = .009$, $p < .05$]. Planned comparisons showed that participants performed better overall when the task was a dual task (92.4%) than when the task was a go–no-go task (89.1%) [$F(1,11) = 7.31$, $p < .05$]. There was no significant difference between the dual task and the

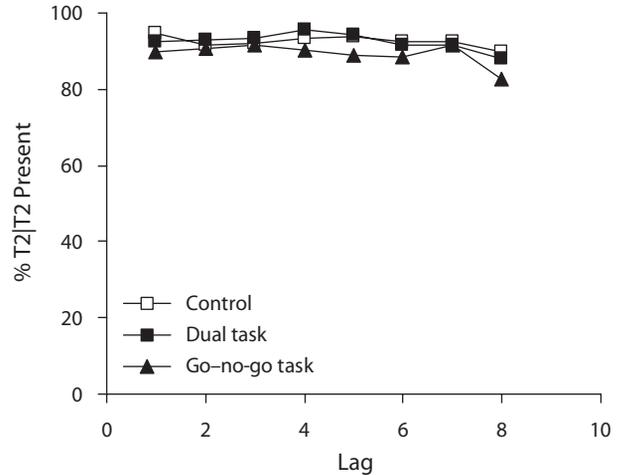


Figure 5. Experiment 4: Mean percentages correct for a visual T2, given that an auditory T1 was correctly identified (when applicable), as a function of lag and task, for those trials on which T1 was a go tone and T2 was present.

control task (92.5%) ($F < 1$). Comparing the go–no-go task with the control task revealed a trend toward a difference [$F(1,11) = 4.66$, $MS_e = .002$, $p = .054$]. The main effect of lag failed to reach significance [$F(7,77) = 2.04$, $MS_e = .011$, $p = .10$]. The two-way interaction between lag and task also failed to reach significance ($F < 1$).

T2 RT Data

Figure 6 presents the mean RT for correct detection of T2 as a function of task and lag for those trials on which T1 was a go tone, T1 was correctly identified and T2 was present.

There was a significant main effect of task and lag [$F(2,22) = 14.60$, $MS_e = 3,018.91$, $p < .001$, and $F(7,77) = 7.63$, $MS_e = 3,198.36$, $p = .001$, respectively]. Of main im-

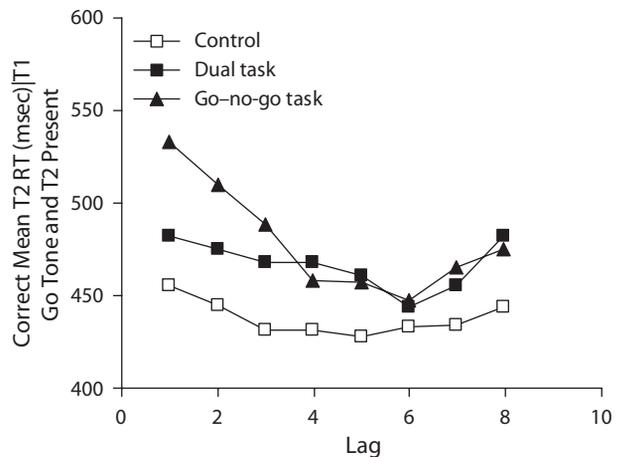


Figure 6. Experiment 4: Mean reaction times (RTs) on correct T2, given that T1 was correctly identified (when applicable), as a function of lag and task, for those trials on which T1 was a go tone and T2 was present.

portance was the significant two-way interaction between task and lag [$F(14,154) = 2.17$, $MS_e = 1,345$, $p < .05$]. Planned comparisons among the control, dual, and go-no-go tasks were performed to clarify the pattern behind this interaction.

Control Versus Dual Task

There was a significant main effect of task [$F(1,11) = 12.51$, $MS_e = 3,268.38$, $p = .005$], since participants were faster overall when the task was a control task (438 msec) than when the task was a dual task (466 msec). Furthermore, the main effect of lag was also significant [$F(7,77) = 3.01$, $MS_e = 1,340.56$], since RTs decreased with increases in lag. The two-way interaction between task and lag failed to reach significance ($F < 1$).

Control Versus Go-No-Go Task

There was a significant main effect of task [$F(1,11) = 27.65$, $MS_e = 3,028.86$, $p < .001$], since participants responded faster when they detected T2 in the control condition (438 msec) than when they detected T2 in the go-no-go task (479 msec). The ANOVA yielded a significant main effect of lag [$F(7,77) = 8.40$, $MS_e = 2,648.29$, $p = .001$], since RTs decreased with lag. Furthermore, the two-way interaction between task and lag was significant [$F(7,77) = 2.74$, $MS_e = 1,541.83$, $p < .05$]. Two-tailed paired samples *t* tests revealed significant differences between the control and go-no-go task for lags 1, 2, 3, 4 (all $ps < .01$), and 7 ($ps < .05$). The differences on other lags failed to reach significance (all $ps > .05$).

Dual Task Versus Go-No-Go Task

Participants were somewhat faster when the task was a dual task (466 msec) than when the task was a go-no-go task (479 msec); however, this main effect of task failed to reach significance [$F(1,11) = 2.75$, $MS_e = 2,759.49$, $p = .13$]. The main effect of lag was significant [$F(7,77) = 6.38$, $MS_e = 3,406.59$, $p = .001$], since RTs decreased with increases in lag. Importantly, the two-way interaction between task and lag was significant [$F(7,77) = 2.37$, $MS_e = 1,365.19$, $p < .05$]. Two-tailed paired samples *t* tests revealed significant differences for lags 1, 2, and 3 ($p = .01$, $p < .05$, and $p < .05$, respectively). The differences on other lags failed to reach significance (all $ps > .4$).

Again, there was no sign of a cross-modal AB. There was no effect of lag on T2 detection accuracy in the dual-task condition. The same was true for the go-no-go task condition: Despite participants now being forced to process T1 before T2, T2 detection accuracy remained as high as in the other conditions, with again no effect of lag.

Direct proof that our manipulation worked (i.e., that T1 was indeed processed before T2) comes from the RTs: In the go-no-go, responses were delayed, indicating that the T2 response had to wait until the T1-related decision was made. The delay appears to reflect a bottleneck on the response decision level, and following PRP logic (Pashler, 1989, 1994), it takes about 400–500 msec after T1 onset before this bottleneck is completely cleared (i.e., lag 4). Note that despite this clear delay and the clear extra de-

mand on T1 processing, T2 accuracy was not affected: The T1 decision did not result in T2's being missed, but only in slower responses. It shows that central, response-related bottlenecks can occur in the absence of an AB and, thus, that these two types of limitations should be dissociated (Arnell & Duncan, 2002).

EXPERIMENT 5 Increased Distractor Similarity Does Not Result in an AB

So far, using simple, stream-like auditory stimuli in combination with visual letters, we have found no sign of a cross-modal AB whatsoever. Experiments 2–4 suggest that echoic memory provides no explanation for this efficient dual tasking. The crucial question then remains as to what causes a cross-modal AB in one case, but not in the other.

A closer look at the literature suggests an alternative hypothesis: An AB occurs when auditory and visual stimuli are of a similar nature (e.g., they are both alphanumeric characters; Arnell & Jenkins, 2004; Arnell & Jolicœur, 1999; Arnell & Larson, 2002). In the case of dissimilar stimuli (e.g., beeps and luminance difference; Hein et al., 2006), the blink appears reduced or absent. Indeed, recent evidence from Arnell and Jenkins (2004) lends some support to this hypothesis. They found that the occurrence of a cross-modal AB depends on the match between the alphanumeric class of T2 and the distractors—that is, both had to be letters. However, their manipulations did not allow them to assess whether the within-modality similarity (i.e., between the visual target and the distractors) is important, or the across-modality similarity (i.e., between the visual T2 and the auditory distractors). Note that in our Experiments 2, 3, and 4, T2 (an “X”) was from the same alphanumeric class as the visual distractor letters, suggesting that the pure visual similarity is not crucial to observing an AB here.

The first possibility we investigated was, therefore, whether the auditory distractors have to be of the same nature as the visual stimuli (including the visual target). For example, when all are letters, an AB may result from the fact that observers first have to ignore letters (and even suppress them, in the auditory domain) and then have to switch to attending to letters (in the visual domain). This switch may induce costs, observed as the AB. Thus, it is not a switch per se that causes the blink (since there was also a task switch in Experiments 2–4), but a switch between very similar, potentially interfering stimulus streams. To test this hypothesis, in the present experiment, we replaced the auditory distractor tones by spoken auditory letters, so that the alphanumeric class of T2 was now from the same alphanumeric class as the auditory distractors. If the visual-target-auditory-distractor similarity plays a crucial role, we would expect to observe a cross-modal AB.

Method

Participants

Twelve new students (6 male; mean age, 20.1 years, ranging from 18 to 22 years) participated in the experiment as paid volunteers.

Apparatus and Stimuli

The present experiment was identical to Experiment 2, except that the distractor tones were replaced by spoken letters. The auditory distractor letters included all the letters of the alphabet except the letters W, N, F, S, and X. Distractor letters were spoken in Dutch in a male voice and were compressed, without altering pitch, to a duration of 90 msec (stereo, 16-bit, 44.1-kHz sample rate) using Cool Edit Pro 2.1 software. They were followed by a 6-msec silent interval, which resulted in an SOA of 96 msec between the different elements in the RSAP stream. The order of the distractor letters in the RSAP stream was randomized.

Results and Discussion

T1 Accuracy

Overall, T1 performance was at 90.0% correct. The ANOVA yielded no significant effect of lag ($F < 1$).

T2 Accuracy

Overall, T2 false alarm rate was at 9.0% in the control condition and at 11.3% in the dual-task condition. Figure 7 presents the mean percentage of correct detection of T2, for those trials on which the target tone (T1) was correctly identified (if applicable) and T2 was present, as a function of lag and task.

There was a trend toward a main effect of task [$F(1,11) = 3.36$, $MS_e = .031$, $p = .094$], reflecting the overall somewhat worse performance in the dual task (69.3%) than in the control task (74.0%). However, the two-way interaction between lag and task failed to reach significance ($F < 1$)—again indicating the absence of a cross-modal AB. The main effect of lag was significant [$F(7,77) = 3.667$, $MS_e = .018$, $p < .01$] but inconsistent with the AB; overall performance decreased, rather than increased, with increases in lag. The better performance at short lags may be explained by the deviating nature of the auditory target (a beep among letters), which may have acted as an alerting or boosting signal for the visual target (Bertelson, 1967; Dalton & Lavie, 2004; Posner & Boies, 1971; Vroomen & De Gelder, 2000).

Thus, despite the fact that the visual target was from the same alphanumeric class as both the visual and the auditory distractors, an AB failed to materialize. This appears to go against Arnell and Jenkins's (2004) claim that the categorical similarity between targets and distractors is crucial to observing a cross-modal AB. In our experiment, observers had little trouble switching from ignoring (auditory) letters to attending to (visual) letters. Note, however, that target–distractor similarity may still be a necessary condition for cross-modal ABs to occur but that it is simply not sufficient when a number of other conditions are not met. Experiment 6 investigates one such condition.

EXPERIMENT 6

A Semantic T1 Generates an AB

In Experiment 5, the auditory distractors were letters, whereas the auditory target (T1) was still a simple beep. This means that even though the distractors were of an alphanumeric nature, the task may not have involved any alphanumeric distinction at all, since observers simply listened out for a beep, regardless of the semantic nature of

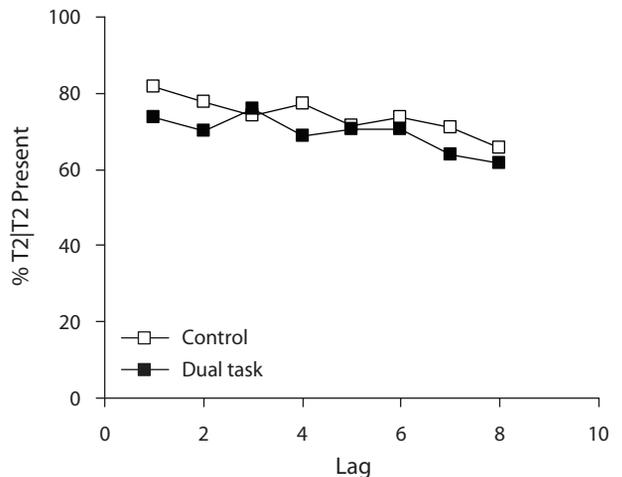


Figure 7. Experiment 5: Mean percentages correct for a visual T2, given that an auditory T1 was correctly identified (when applicable) and T2 was present, as a function of lag and task.

the distractors. Hence, since there was no semantic analysis of the auditory stream, a cross-modal switch to a semantic analysis of the visual stream would incur little cost. In other words, a cross-modal AB may be observed when both the auditory and the visual streams demand semantic analyses—analyses that may then interfere with each other, resulting in competition for resources or switch costs.

To test this hypothesis, we replaced the T1 beep with a digit. This ensured that observers had to apply a semantic analysis to the auditory stream. Such semantic analysis may interfere directly with the semantic analysis necessary for the visual stream. For example, having to distinguish digits from letters in the auditory stream may result in the suppression of letters. This leads to a cost when observers then need to switch to analyzing letters for a visual target.

Method

Participants

Ten new students (4 male; mean age, 21.7 years, ranging from 18 to 27 years) participated in the experiment as paid volunteers.

Apparatus and Stimuli

The present experiment was identical to Experiment 5, except that the target tone (T1) was replaced by one of the spoken digits 1, 2, 3, and 4. Overall, the physical properties of the digits were identical to those of the distractor letters.

Design and Procedure

Participants were asked to make an unspeeded response to T1 by pressing the 1, 2, 3, or 4 key when they identified the digit 1, 2, 3, or 4, respectively, and to make an unspeeded response to T2 by pressing the n or j key when the letter “X” was absent or present, respectively. Prior to the experiment, participants heard the digits in order to get familiar with the target.

Results and Discussion

T1 Accuracy

Overall, T1 performance was at 95.5% correct. The ANOVA yielded no significant effect of lag ($F < 1$).

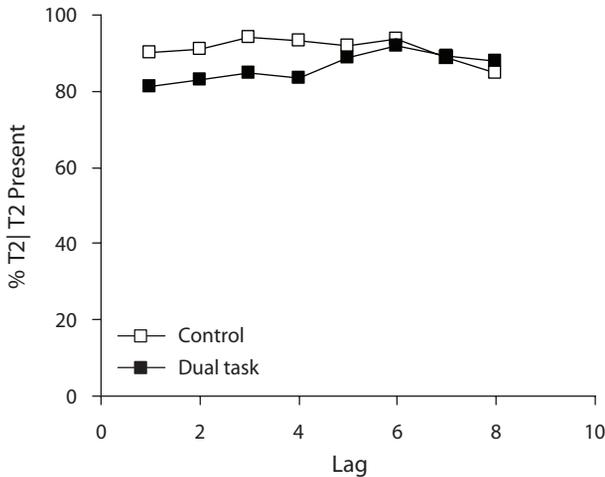


Figure 8. Experiment 6: Mean percentages correct for a visual T2, given that an auditory T1 was correctly identified (when applicable) and T2 was present, as a function of lag and task.

T2 Accuracy

Overall, T2 false alarm rate was at 5.8% in the control condition and at 6.9% in the dual-task condition. Figure 8 presents the mean percentages of correct detection of T2, for those trials on which the target tone (T1) was correctly identified (when applicable) and T2 was present, as a function of lag and task.

There was a significant main effect of task [$F(1,11) = 8.27$, $MS_e = .015$, $p < .05$], since participants performed better in the control condition (91.5%) than in the dual-task condition (86.4%). The main effect of lag reached significance [$F(7,77) = 2.60$, $MS_e = .005$, $p < .05$], since overall performance increased with increases in lag. Importantly, the two-way interaction between task and lag was significant [$F(7,77) = 2.82$, $MS_e = .016$, $p < .02$], suggesting a cross-modal AB. This interaction was further analyzed in detail by pairwise t tests for each lag. These revealed significant differences between the dual-task and the control conditions at lags 1 ($p < .005$), 2, 3, and 4 (all $ps < .05$). The differences at other lags failed to reach significance (all $ps > .25$).

Within the present series of experiments, the observed effects constitute the first sign of a cross-modal AB, consistent with many other studies (Arnell & Jenkins, 2004; Arnell & Jolicœur, 1999; Arnell & Larson, 2002; Potter et al., 1998). Contrast this with our earlier experiments, in which there was no sign of an AB whatsoever. The results suggest that, in order to observe a cross-modal AB, the auditory and visual tasks must compete on a similar level of representation. In the present case, this competition played out on a semantic level: Extracting an auditory digit from letters interfered with extracting a visual letter from visual letters. It may be the case that a cross-modal AB is more easily observed with semantic stimuli exactly because of this level playing field between audition and vision. Other, similar auditory and visual tasks that are not semantic in nature may be possible but are somewhat harder to imagine (e.g., what would be the visual equiva-

lent to a beep?). We will return to this in the General Discussion section.

Finally, note that the emergence of an AB in the present case was not simply due to the T1 task's becoming more difficult (i.e., extracting a digit instead of a beep). If anything, T1 performance (95.5%) was better than in Experiment 5 (90.0%) [$t(22) = 2.41$, $p < .05$].

GENERAL DISCUSSION

Whereas Experiment 1 revealed clear AB patterns within the auditory and visual domains, Experiments 2, 3, and 4 failed to find similar patterns across the two modalities, even though the respective auditory and visual stimuli were exactly the same as those used in Experiment 1. The results corroborate earlier failures to find a cross-modal AB (Duncan et al., 1997; Hein et al., 2006; Soto-Faraco & Spence, 2002).

An important contribution of the present study is the exclusion of a previously unexplored explanation for the lack of cross-modal interference in the AB paradigm. We hypothesized that observers may be able to postpone auditory processing until after the visual task has been completed by making use of echoic memory (cf. Arnell, 2006; Chun & Potter, 2001). Such a strategy may be especially helpful when relatively simple auditory stimuli are used, such as streams of beeps, with the target being a deviation from this stream (as in the present study; Duncan et al., 1997; Hein et al., 2006). However, the present results strongly argue against this hypothesis. Under the assumption that sounds would interfere with retrieval (Deutsch, 1970), in Experiment 3 the use of echoic memory was discouraged by presenting distracting sounds during the response stage. This did not lead to a different pattern of results. In Experiment 4, postponement of auditory T1 processing was completely prevented by making T1 directly relevant for the (speeded) T2 response. It was found that T1 affected T2 performance in terms of RTs (since T2 responses were delayed), but not in terms of accuracy (i.e., T2 was missed no more often). In other words, again, there was no sign of an AB. Taken together, the results show that postponement of auditory target processing is an unlikely explanation for the absence of a cross-modal AB.

The important question then remained as to what is the explanation for the absence of a cross-modal AB in some cases (Experiments 2–4; Duncan et al., 1997; Hein et al., 2006; Soto-Faraco & Spence, 2002), but not in others (Arnell & Jenkins, 2004; Arnell & Jolicœur, 1999; Potter et al., 1998). An important earlier hypothesis has been that cross-modal deficits are observed when a task switch is involved (e.g., from detecting digits to detecting letters), but that the AB proper is a modality-specific deficit (Potter et al., 1998). However, in the present study, there should have been a clear task switch between the auditory task (identifying a beep) and the visual task (detecting an "X"), yet there was no cross-modal deficit (see also Hein et al., 2006, for a task switch without a cross-modal AB). Furthermore, one may argue that, by definition, any cross-modal paradigm involves a task switch. After all, the observer

needs to switch from the auditory task to the visual task (or whichever modalities are involved). Moreover, Arnell and Larson (2002), as well as Arnell and Jenkins (2004), succeeded in observing a cross-modal AB even when no task switch was involved (in terms of to-be-reported target categories; e.g., both T1 and T2 were letters).

Experiments 5 and 6 shed more light on what may underlie cross-modal deficits. Arnell and Jenkins (2004) proposed that the occurrence of a cross-modal AB depends on the match between the alphanumeric class of T2 and the distractors. When all distractors (both visual and auditory) were letters, they observed clear cross-modal ABs when T2 was a letter, but no AB when T2 was a digit, suggesting an important role for target–distractor similarity. Within the visual domain, the contribution of target–distractor similarity to the AB has been recognized before (Chun & Potter, 1995; Isaak, Shapiro, & Martin, 1999; Maki, Frigen, & Paulson, 1997; Raymond, Shapiro, & Arnell, 1995). These similarity effects have usually been explained in terms of more effective masking and/or increased competition between targets and distractors (i.e., T2 is more effectively masked by, and suffers more from competition from, the trailing distractor when this distractor is very similar; see, e.g., Shapiro, Arnell, & Raymond, 1997). However, such visual effects could not account for the present data since, although the visual T2 was from the same alphanumeric class as the visual distractors and the visual T2 task was identical in the unimodal (Experiment 1) and cross-modal (Experiments 2–4) cases, an AB was observed in the one, but not in the other. Experiment 5 therefore tested for the possibility that the visual T2 had to be of the same class as the auditory distractors, rather than the visual distractors. However, replacing the auditory beeps with spoken letters did not result in a cross-modal AB.

Instead, Experiment 6 suggests that, in addition to using similar distractors (as shown by Arnell & Jenkins, 2004), the T1 and T2 *tasks* also have to be similar. For example, as manipulated here, both tasks involve a semantic analysis of their respective streams. The semantic analysis of an auditory target may take resources away from a similar analysis of a visual T2. Alternatively, the semantic analysis of the auditory target may result in the semantic suppression of the auditory distractors, which then carries over to a visual target of the same semantic nature. Within the visual domain, it is known that semantic processing affects the magnitude of the AB (see, e.g., Davenport & Potter, 2005; Maki et al., 1997; Potter, Dell'Acqua, Pesciarelli, Job, & Peressotti, 2005). From the cross-modal literature, it also appears that auditory–visual ABs are more robust when semantic stimuli are involved (e.g., Arnell & Jenkins, 2004; Arnell & Jolicœur, 1999; Potter et al., 1998) than when lower level, meaningless beeps or syllables are used (e.g., Duncan et al., 1997; Hein et al., 2006). This is probably not because meaningless stimuli cannot, in principle, generate an AB, but because it is difficult to create conditions of nonsemantic but, nonetheless, very similar auditory and visual stimuli. Semantic processes are one of the few that the two have in common. Another dimension that the two modalities have in common is space. The prediction would then be that auditory and visual tasks interfere on a spatial,

rather than semantic, level. Such interference was indeed observed by Theeuwes, Van der Burg, Olivers, and Bronkhorst (2006, Experiment 4). It would be interesting to see whether cross-modal interference also emerges when other common processes or representations are invoked, such as those involved in object recognition. For example, would the task of classifying different animal sounds interfere with the recognition of animal pictures?

Whereas the different senses appear to require a common playground in order to interfere with each other, such a level playing field is not necessary within modalities. Second-target deficits have been observed even when the first and the second targets involved completely different tasks, sometimes requiring little semantic processing (e.g., Joseph, Chun, & Nakayama, 1997; Ross & Jolicœur, 1999), probably because, even when T1 and T2 involve different tasks, if the stimuli are derived from the same modality, these tasks are still likely to tap into similar lower level processes, such as those involved in visual and auditory feature extraction and object recognition. Furthermore, whereas it may be difficult to create an AB paradigm without a task-switching component between modalities (since a modality switch may automatically imply a task switch), within modalities it is probably very important to distinguish between paradigms that do and those that do not involve a task switch (Chun & Potter, 2001), because task switches put a direct strain on attentional resources, aggravating any AB. Recent studies have even questioned whether the “attentional blink proper” (i.e., without a task switch) is caused at all by competition for and depletion of such attentional resources (Di Lollo, Kawahara, Ghorashi, & Enns, 2005; Olivers, in press; Olivers, Van der Stigchel, & Hulleman, 2007; Raymond et al., 1992). Instead of T1’s using up resources that are then no longer available to T2, these studies assign a much more important role to the post-T1 distractors. These distractors trigger an attentional state in which similar input is no longer accepted (e.g., the gate to consciousness is closed to prevent further interference), causing an AB. It remains to be seen to what extent such distractor-driven processes also play a role in the cross-modal AB. The fact that the relationship between the targets and the distractors appears important (Arnell & Jenkins, 2004; Experiment 6 here) provides one indication that they do.

CONCLUSION

The typical difficulty in obtaining an auditory–visual AB is not due to observers’ making strategic use of echoic memory. Instead, auditory–visual interference appears to depend on shared representations between the modalities. Here, we found cross-modal interference on a semantic level, but interference may also be found for other shared representations, such as those involving spatial coding and object recognition.

AUTHOR NOTE

For this research, J.T. and A.W.B. were supported by Dutch Technology Foundation STW Grant VET.7079, by the applied science division of NWO, and by the Technology Program of the Ministry of Economic Affairs; C.N.L.O. was supported by an NWO-VENI grant. Correspondence

concerning this article should be addressed to E. Van der Burg, Department of Cognitive Psychology, Vrije Universiteit, Boechorststraat 1, 1081BT Amsterdam, The Netherlands (e-mail: e.van.der.burg@psy.vu.nl).

REFERENCES

- ARNELL, K. M. (2006). Visual, auditory, and cross-modality dual-task costs: Electrophysiological evidence for an amodal bottleneck on working memory consolidation. *Perception & Psychophysics*, **68**, 447-457.
- 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., & JENKINS, R. (2004). Revisiting within-modality and cross-modality attentional blinks: Effects of target-distractor similarity. *Perception & Psychophysics*, **66**, 1147-1161.
- 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.
- ARNELL, K. M., & LARSON, J. M. (2002). Cross-modality attentional blink without preparatory task-set switching. *Psychonomic Bulletin & Review*, **9**, 497-506.
- BERTELSON, P. (1967). The time course of preparation. *Quarterly Journal of Experimental Psychology*, **19**, 272-279.
- BREGMAN, A. S. (1978). Auditory streaming is cumulative. *Journal of Experimental Psychology: Human Perception & Performance*, **4**, 380-387.
- CHUN, M. M., & POTTER, M. C. (1995). A two-stage model for multiple detection in RSVP. *Journal of Experimental Psychology: Human Perception & Performance*, **21**, 109-127.
- CHUN, M. M., & POTTER, M. C. (2001). The attentional blink and task switching within and across modalities. In K. Shapiro (Ed.), *The limits of attention: Temporal constraints in human information processing* (pp. 20-35). London: Oxford University Press.
- COWAN, N. (1984). On short and long auditory stores. *Psychological Bulletin*, **96**, 341-370.
- CROWDER, R. G. (1993). Auditory memory. In S. McAdams & E. Bigand (Eds.), *Thinking in sound: The cognitive psychology of human audition* (pp. 113-145). Oxford: Oxford University Press.
- DALTON, P., & LAVIE, N. (2004). Auditory attentional capture: Effects of singleton distractor sounds. *Journal of Experimental Psychology: Human Perception & Performance*, **30**, 180-193.
- DAVENPORT, J. L., & POTTER, M. C. (2005). The locus of semantic priming in RSVP target search. *Memory & Cognition*, **33**, 241-248.
- DELL'ACQUA, R., JOLICŒUR, P., PESCIARELLI, F., JOB, R., & PALOMBA, D. (2003). Electrophysiological evidence of visual encoding deficits in a cross-modal attentional blink paradigm. *Psychophysiology*, **40**, 629-639.
- DEUTSCH, D. (1970). Tones and numbers: Specificity of interference in short-term memory. *Science*, **168**, 1604-1605.
- DI LOLLO, V., KAWAHARA, J., GHORASHI, S. M. S., & ENNS, J. T. (2005). The attentional blink: Resource depletion or temporary loss of control? *Psychological Research*, **69**, 191-200.
- DORMAN, M. F., CUTTING, J. E., & RAPHAEL, L. J. (1975). Perception of temporal order in vowel sequences with and without formant transitions. *Journal of Experimental Psychology: Human Perception & Performance*, **1**, 121-129.
- DUNCAN, J., MARTENS, S., & WARD, R. (1997). Restricted attentional capacity within but not between sensory modalities. *Nature*, **387**, 808-810.
- HEIN, G., PARR, A., & DUNCAN, J. (2006). Within modality and cross-modality attentional blinks in a simple discrimination task. *Perception & Psychophysics*, **68**, 54-61.
- HILLSTROM, A., SHAPIRO, K. L., & SPENCE, C. (2002). Attentional limitations in processing sequentially presented vibrotactile targets. *Perception & Psychophysics*, **64**, 1068-1081.
- 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.
- JOSEPH, J. S., CHUN, M. M., & NAKAYAMA, K. (1997). Attentional requirements in a "preattentive" feature search task. *Nature*, **387**, 805-807.
- KALLMAN, H. J., & MASSARO, D. W. (1979). Similarity effects in backward recognition masking. *Journal of Experimental Psychology: Human Perception & Performance*, **5**, 110-128.
- MAKI, W. S., FRIGEN, K., & PAULSON, K. (1997). Associative priming by targets and distractors during rapid serial visual presentation: Does word meaning survive the attentional blink? *Journal of Experimental Psychology: Human Perception & Performance*, **23**, 1014-1034.
- OLIVERS, C. N. L. (2007). The time course of attention: It is better than we thought. *Current Directions in Psychological Science*, **16**, 11-15.
- OLIVERS, C. N. L., VAN DER STIGCHEL, S., & HULLEMAN, J. (2007). Spreading the sparing: Against a limited-capacity account of the attentional blink. *Psychological Research*, **71**, 126-139.
- PASHLER, H. (1989). Dissociations and dependencies between speed and accuracy: Evidence for a two-component theory of divided attention in simple tasks. *Cognitive Psychology*, **21**, 469-514.
- PASHLER, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin*, **116**, 220-244.
- POSNER, M. I., & BOIES, S. J. (1971). Components of attention. *Psychological Review*, **78**, 391-408.
- POTTER, M. C., CHUN, M. M., BANKS, B. S., & MUCKENHOPT, 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.
- POTTER, M. C., DELL'ACQUA, R., PESCIARELLI, F., JOB, R., & PERESSOTTI, F. (2005). Bidirectional semantic priming in the attentional blink. *Psychonomic Bulletin & Review*, **12**, 460-465.
- 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.
- ROSS, N. E., & JOLICŒUR, P. (1999). Attentional blink for color. *Journal of Experimental Psychology: Human Perception & Performance*, **25**, 1483-1494.
- SHAPIRO, K. L., ARNELL, K. M., & RAYMOND, J. E. (1997). The attentional blink. *Trends in Cognitive Sciences*, **1**, 291-296.
- SHAPIRO, K. L., RAYMOND, J. E., & ARNELL, K. M. (1994). Attention to visual pattern information produces the attentional blink in RSVP. *Journal of Experimental Psychology: Human Perception & Performance*, **20**, 357-371.
- SHEN, D., & MONDOR, T. A. (2006). Effect of distractor sounds on the auditory attentional blink. *Perception & Psychophysics*, **68**, 228-243.
- SOTO-FARACO, S., & SPENCE, C. (2002). Modality-specific auditory and visual temporal processing deficits. *Quarterly Journal of Experimental Psychology*, **55A**, 23-40.
- SOTO-FARACO, S., SPENCE, C., FAIRBANK, K., KINGSTONE, A., HILLSTROM, A., & SHAPIRO, K. (2002). A cross-modal attentional blink between vision and touch. *Psychonomic Bulletin & Review*, **9**, 731-738.
- THEEUWES, J., VAN DER BURG, E., OLIVERS, C. N. L., & BRONKHORST, A. W. (2006). Cross-modal interactions between sensory modalities: Implications for the design of multisensory displays. In A. F. Kramer, D. A. Wiegmann, & A. Kirlik (Eds.), *Attention: From theory to practice* (pp. 196-208). Oxford: Oxford University Press.
- TREMBLAY, S., VACHON, F., & JONES, D. M. (2005). Attentional and perceptual sources of the auditory attentional blink. *Perception & Psychophysics*, **67**, 195-208.
- VACHON, F., & TREMBLAY, S. (2005). Auditory attentional blink: Masking the second target is necessary, delayed masking is sufficient. *Canadian Journal of Experimental Psychology*, **59**, 279-286.
- VROOMEN, J., & DE GELDER, B. (2000). Sound enhances visual perception: Cross-modal effects of auditory organization on vision. *Journal of Experimental Psychology: Human Perception & Performance*, **26**, 1583-1590.