Journal of Experimental Psychology:
Learning, Memory, and Cognition
1998, Vol. 24, No. 4, 979–992

0278-7393/98/$3.00

Two Attentional Deficits in Serial Target Search: The Visual Attentional Blink and an Amodal Task-Switch Deficit

Mary C. Potter
Massachusetts Institute of Technology

Marvin M. Chun
Yale University

Bradley S. Banks and Margaret Muckenhoupt
Massachusetts Institute of Technology

When monitoring a rapid serial visual presentation at 100 ms per item for 2 targets among distractors, viewers have difficulty reporting the 2nd target (T2) when it appears 200–500 ms after the onset of the 1st letter target (T1): an attentional blink (AB; M. M. Chun & M. C. Potter, 1995b; J. E. Raymond, K. L. Shapiro, & K. M. Arnell, 1992). Does the same deficit occur with auditory search? The authors compared search for auditory, visual, and cross-modal targets in 2 tasks: (a) identifying 2 target letters among digits (Experiments 1–3 and 5) or digits among letters (Experiment 6), and (b) identifying 1 digit among letters and deciding whether an *X* occurred among the subsequent letters (Experiment 4). In the experiments using the 1st task, the standard AB was found only when both targets were visual. In the 2nd task, with a change in selective set from T1 to T2, a task-switching deficit was obtained regardless of target modality.

When participants search for two targets among distractors in a rapidly presented, serial list, the second target (T2) is often missed if it falls in a window 200–500 ms after the onset of the first target (Broadbent & Broadbent, 1987; Chun & Potter, 1995b; Raymond, Shapiro, & Arnell, 1992; Weichselgartner & Sperling, 1987). This deficit has been termed the *attentional blink* (AB) by Raymond et al. (1992). A typical result is shown in Figure 1 (Chun & Potter, 1995b). In this experiment participants searched for letters appearing among digit distractors, with each stimulus presented for 100 ms with no interstimulus interval (ISI). The first letter, T1, was reported on more than 80% of the trials. The figure shows the probability of reporting T2, given that T1 was correctly reported, at each of eight lags. Note two findings that are characteristic of this task: When the two targets are adjacent (Lag 1), there is little or no deficit, but at Lag 2 (with just one intervening digit distractor) performance on T2 drops sharply, and recovers gradually over the next 300–400 ms. At longer lags, performance on the second target is not much lower than on the first target.

Chun and Potter (1995b) proposed a two-stage model of the AB deficit in which they posited that at the rate of 100 ms per item all or most items are identified (Stage 1) as they appear, but they are rapidly overwritten and forgotten unless they are selected for further processing in a serial second stage. The difficulty of encoding T1 in Stage 2 determines the presence and time course of the deficit in reporting T2. Encountering a letter target triggers this second stage, but attentional selection is imprecise and typically the input to the second stage includes not only the target letter, but also the next item in the sequence. Hence, in Stage 2 the viewer has to discriminate between the target and the following item, and encode the target in a durable short-term memory. When the following item is itself a target, then the two targets will be processed together, and both are likely to be reported. (See Chun & Potter, 1995b, for evidence that the two items are processed together, but with some competition between them.) The second stage is strictly serial, so no further items can enter it until the current processing is completed. In consequence, a second target that arrives while the first target is still being processed in Stage 2 will be identified fleetingly but is frequently overwritten by subsequent items (see Giesbrecht & Di Lollo, in press) before it can be processed in Stage 2.

This model reflects properties of the visual system, which has only a short-lasting iconic buffer to maintain information that is not being attended to. We proposed that the first stage is a postcategorical buffer (e.g., Potter, 1993) rather than an early visual buffer of the kind originally termed *iconic*. (We address below the question of whether the postcategorical

Mary C. Potter, Bradley S. Banks, and Margaret Muckenhoupt, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology; Marvin M. Chun, Department of Psychology, Yale University. Margaret Muckenhoupt is now in independent practice in Somerville, Massachusetts.

Parts of this research were reported at the 36th and 37th annual meetings of the Psychonomic Society in November 1995 in Los Angeles, California, and in November 1996 in Chicago, Illinois. This research was supported by Grant MH47432 from the National Institute of Mental Health, by Grant BNS90-13026 from the National Science Foundation, and by Postdoctoral Grant EY06592 from the National Institutes of Health. We thank Chris Hooker, Kevin Doyle, Marsha Novak, Susan Rushing, Margo Harbaugh, Manjari Chanda, Sarah Elovich, Tim Chklovski, and Janina Rado for their assistance with the research.

Correspondence concerning this article should be addressed to Mary C. Potter, Department of Brain and Cognitive Sciences, E10-039, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139. Electronic mail may be sent to potter@psyche.mit.edu.

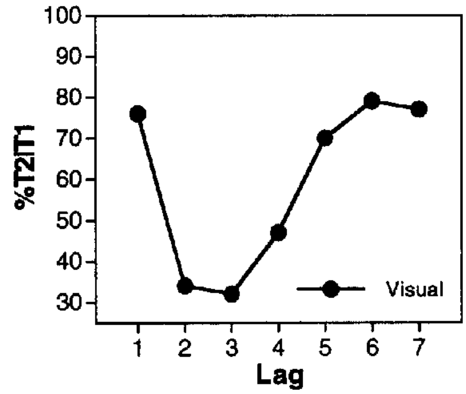

Figure 1. The attentional blink effect: correct report of the second target (T2) given correct report of the first target (T1) as a function of lag; the stimulus onset asynchrony was 100 ms. From "A Two-Stage Model for Multiple Target Detection in Rapid Serial Visual Presentation," by M. M. Chun and M. C. Potter, 1995, *Journal of Experimental Psychology: Human Perception and Performance, 21,* p. 112. Copyright 1995 by the American Psychological Association. Adapted with permission.

buffer is specific to the visual system or is amodal.) The assumption that the initial processing results in a postcategorical representation is also inherent in other late-selection models of AB (Shapiro & Raymond, 1994). Neurophysiological support as well as behavioral evidence for semantic activation of blinked items exists (Luck, Vogel, & Shapiro, 1996; Maki, Frigen, & Paulson, 1997; Shapiro, Driver, Ward, & Sorensen, 1997), supporting the assumption that Stage 1 processing results in a postcategorical representation.

Now consider what one might expect in an auditory version of the AB procedure. If the first-stage, preattentive representation is echoic one might not expect the same acute sensitivity to the lag between the two targets (see Crowder, 1993, for a review of the auditory memory literature), inasmuch as the first-stage representation of T2 would still be available when the listener has finished processing T1, even with a short lag. But if the first-stage representation is at a postcategorical stage that is the same for vision and audition—for example, an abstract representation of the identity of each letter and digit—then one might expect a similar attentional deficit for the auditory and visual modalities.

Another possibility is that there are multiple bottlenecks in the processing stream. There may be modality-specific bottlenecks for rapidly presented visual and auditory information and also there may exist a general bottleneck that operates on amodal representations. By examining the types of processing interference that are triggered by target detection both within and across the visual and auditory modalities, we can begin to isolate the nature and architecture of critical bottlenecks within the human information processing stream.[1]

Here we report six experiments in which auditory and visual target search were compared. In the first two experiments we compared visual and auditory attentional deficits, and in a third experiment we looked at possible crossmodal attentional deficits when the first target was auditory and the second visual, or vice versa. In the fourth experiment we examined both within-modality and cross-modal conditions in a procedure that required a switch in selective set between the first and the second targets. The fifth and sixth experiments followed up those results. In all six experiments the stimuli were written or spoken letters of the alphabet or arabic numerals.

For the present experiments we developed compressed-speech versions of a set of letters and digits. In Experiments 1 and 3 we adjusted the duration of the auditory stimuli and the contrast of the visual stimuli to match performance on detection of the first target, T1: The stimulus onset asynchrony (SOA) was 135 ms in the auditory condition and 120 ms in the visual condition. In Experiment 2 we matched the SOAs of the visual and auditory stimuli: Both were 135 ms. In Experiments 4–6 the SOAs for both visual and auditory stimuli were 120 ms.

Experiment 1

Method

Participants. Twenty-four observers from the Massachusetts Institute of Technology's paid volunteer pool participated in Experiment 1. An additional 2 participants were replaced because they did not meet a performance criterion (see the *Scoring* section).

Design and procedure. The stimulus set consisted of 17 letters and six digits, excluding 9 letters (*A, G, H, I, L, O, Q, V, W*) and four digits (0, 1, 7, 8) because they were either visually or auditorily difficult to identify or to discriminate from other letters or digits. The auditory stimuli were individually recorded in a female voice, compressed, and digitized for presentation by the computer. The duration of each of the auditory stimuli was 120 ms or less (averaging 113 ms), and silence was added to bring each stimulus to 120 ms. For comprehensibility, a further 15-ms silent interval was added to each stimulus, for a total frame duration of 135 ms. The visual stimuli consisted of uppercase letters and arabic digits, presented for 120 ms each in the center of the screen.

Each trial consisted of one or two single-letter targets presented among digit distractors, for a total of 13 characters. The letter targets and distractors were drawn randomly from the 18 letters and

[1] Two earlier unpublished reports in which rapid auditory presentation was used in a two-target search task should be noted. Kanwisher (1994) reported a study in which Lag 1 and Lag 4 conditions could be compared in three two-target experiments, with auditory or visual stimuli: The data suggested AB for visual targets but little or none for auditory targets. Because these were incidental observations in an experiment addressed to other questions, Kanwisher did not test the significance of the visual and auditory difference. Shulman (1994) reported an attentional deficit for the second of two auditory targets, although he found some differences between the auditory and visual patterns. Shulman, however, had other criticisms of both the visual and auditory AB findings, which he attributed to interference from subsequent distractor stimuli. To address Shulman's main criticism, we (Chun & Potter, 1995a), carried out a visual AB experiment comparing trials in which a negative correlation of lag with the number of distractors following T2 was present or absent (in the latter condition there were six distractors following T2, regardless of lag between T1 and T2). The two types of trials were intermixed randomly. We obtained an AB effect in both conditions, and there was no interaction between lag and condition.

six digits, respectively, with the constraint that T1 and T2 were never the same letter, and digit distractors were not repeated within a lag of four characters. Participants were divided into two groups of 12, one the experimental group (with two targets per trial) and one the control group (with only the second target per trial). Note that we did not use the control condition of Raymond et al. (1992) in which the participant was simply instructed to ignore T1 and report T2. In pilot work Chun (1992) had found that a participant cannot readily perform this task when targets are all in one category and are not marked by a distinctive cue (such as being white rather than black). This difficulty is presumably related to differences in the T1 and T2 tasks and the task-switch issue, and is taken up in the General Discussion.

For the experimental group there were three counterbalanced, within-subject variables: stimulus modality (visual or auditory), which was blocked; serial position of T1 (Positions 3, 4, or 5); and lag between T1 and T2 (Lags 1–7, with zero to six intervening digit distractors). Each modality block included 10 replications of the 21 T1 × Lag conditions, randomized in smaller blocks of 42 trials. The order of the two modalities was counterbalanced between subjects. The choice of targets and distractors and the order of conditions was separately randomized for each participant. The control group had the same sequences as the experimental group, except that T1 was replaced by a distractor digit. Thus, lag was a dummy variable, corresponding to the serial position of T2. Because the T1 serial position in the experimental group varied from 3–5, the actual serial position of T2 varied equivalently, at each lag.

During both the visual and auditory blocks participants wore headphones. They were instructed to listen for or look for two letter targets (or one target, in the control condition) among the digits, and when the sequence ended they typed the letters, in order if possible, on the keyboard. They were encouraged to guess if uncertain, but to press the space bar instead if they had no idea about the identity of one or both letters. Participants initiated each trial by pressing the space bar, and the trial began after a brief delay. On visual trials there was a fixation point that preceded the stimuli by 390 ms, and an ampersand followed the last stimulus. Auditory trials began shortly after the participant pressed the space bar; there was no initial or final cue.

Apparatus. The experiment was conducted using a Macintosh Quadra 840 AV equipped with a 14-in. Macintosh Color Display Monitor and an Audiomedia II Sound Card, Model MM005. MacProbe version 1.6.9 by Aristometrics (Hunt, 1994) was used to program and administer the task. The background color for the monitor was set at 127, a standard Macintosh color on a gray scale where 0 equals black and 225 equals white. The fixation plus sign was color 200 and the visual characters (digits and letters) appeared in color 140 in 24-point Monaco type.

Scoring. Two participants, 1 in the experimental and 1 in the control group, were replaced because they failed to meet a criterion of 64% of trials on which T1 was reported correctly, calculated separately for the visual and the auditory conditions. Targets were scored as correct regardless of the order in which they were reported (Chun & Potter, 1995b).

Results and Discussion

In the experimental group, T1 was reported correctly on 87.4% of the trials in the visual condition and 86.7% of the trials in the auditory condition ($F < 1$). The probability of reporting T2 given that T1 had been reported (T2|T1), is shown for the visual and auditory conditions in Figure 2. An analysis of variance (ANOVA) with modality and lag as within-subject variables showed no main effect of modality: The means for visual and auditory conditions were both .78 ($F < 1$). There was a significant main effect of lag, $F(6, 66) = 8.14, p < .001$, and a significant interaction between lag and modality, $F(6, 66) = 6.10, p < .001$.

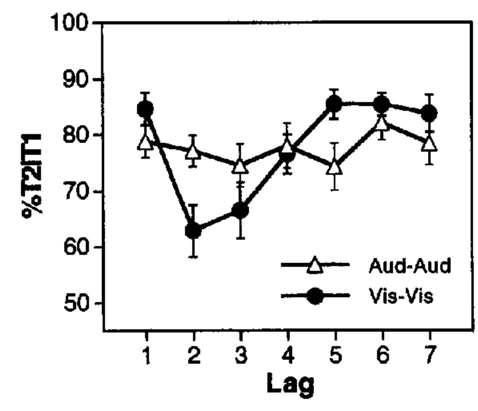

Figure 2. Correct report of the second target (T2), given correct report of the first target (T1) as a function of lag, for visual (Vis) stimuli (stimulus onset asynchrony [SOA] = 120 ms) and auditory (Aud) stimuli (SOA = 135 ms) in Experiment 1. Vertical bars show standard error of the mean.

Figure 2 shows the interaction. There is the usual AB pattern for the visual condition: good performance at Lag 1, and a marked drop at Lags 2 and 3, with full recovery at Lag 5. In contrast, with auditory stimuli there is no evidence for a lag effect.

To assess the presence of a deficit for the second target, we compared the unconditional report of T2 in the experimental group to that of the control group, who only saw T2. The results for the two visual conditions are shown in Figure 3A. An ANOVA of the two visual conditions showed a significant difference between the experimental and control groups, $F(1, 22) = 16.86, p < .001$. There was a main effect of lag, $F(6, 132) = 10.36, p < .001$, and a significant interaction between group and lag, $F(6, 132) = 12.14, p < .001$. The experimental group showed the marked AB effect found in the conditional analysis; the control group showed no lag effect.

The results for the two auditory conditions are shown in Figure 3B. An ANOVA showed a main effect of group, $F(1, 22) = 22.46, p < .001$, indicating that T2 was less likely to be reported when T1 was presented. There was no main effect of lag, $F(6, 132) = 1.09, p > .37$. The interaction between group and lag was marginally significant, however, $F(6, 132) = 2.09, p < .06$. In the control condition in which only T2 was shown and reported, there were no significant effects of lag. As in the conditional analysis, there was little evidence for a systematic lag effect for auditory stimuli in the experimental group. Thus, although there was a substantial overall difference between the experimental and control conditions, showing that it was more difficult to report T2 if T1 had to be attended to and reported, there was no apparent lag effect for T2 in the auditory condition.

Comparing T2 performance in the visual and auditory

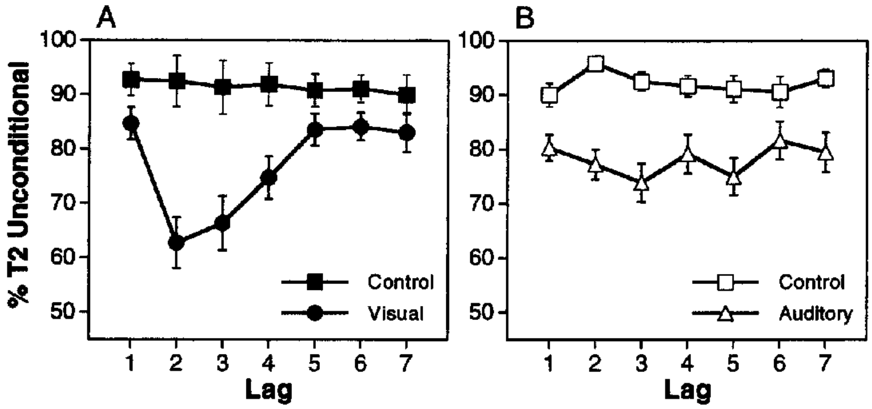

Figure 3. Correct report of the second target (T2; unconditional) as a function of lag, for the experimental group and the T2-only control group in Experiment 1, in (A) the visual condition and (B) the auditory condition. Vertical bars show standard error of the mean.

conditions, it is clear that only the visual condition shows the characteristic AB effect, although there is a net T2 deficit of similar average magnitude (across all lags) for both modalities. There was a highly significant interaction between group (experimental and control) and lag in the visual case, but only a marginal interaction in the auditory case. These differences in the patterns of interference with visual and auditory targets occurred even though the two modalities were matched in difficulty (as measured by T1 detection performance). These results indicate that attention can be deployed differently when the targets are auditory than when they are visual: Auditory attention apparently does not blink.

Experiment 2

Experiment 2 replicated Experiment 1's experimental condition, except that the rates of presentation in the two modalities were matched.

Method

Participants. Twelve individuals from the same pool as that of Experiment 1 participated. None had been in previous AB experiments.

Design and procedure. Because the actual rate of presentation for the visual and auditory conditions differed by 15 ms in Experiment 1, we repeated the experiment with matched durations. To match the durations, each visual letter or digit was followed by a 15-ms blank interval, bringing the SOA to 135 ms. The auditory stimuli were the same as in Experiment 1, with an SOA of 135 ms that included a silent interval of 15 ms or somewhat more, depending on the precise duration of the auditory stimuli. In all other respects the method was the same as that of Experiment 1, except that there was no one-target control group.

Results and Discussion

In this experiment T1 was reported correctly on 94.5% of the visual trials and 86.4% of the auditory trials, $F(1, 11) = 8.04$, $p < .05$. The probability of reporting T2, given correct report of T1, is shown for both visual and auditory modalities in Figure 4. An ANOVA showed no significant main effect of modality ($F < 1$); a significant effect of lag; $F(6, 66) = 5.14$, $p < .001$; and a significant interaction between lag and modality, $F(6, 66) = 3.04$, $p < .05$: the same pattern of results as in Experiment 1's experimental group. An analysis of Experiments 1 and 2 together showed no significant main effect of experiment ($F = 1.31$) and no interactions with experiment. The interaction between lag and modality was again highly significant, $F(6, 132) = 7.84$, $p < .001$.

Thus, the overall results of Experiments 1 and 2 suggest that there is some attentional deficit for the second of two targets whether they are visual or auditory, but the temporal pattern is very different in the two modalities: In the visual condition there is a distinct blink, but in the auditory condition the deficit is spread across all lags, irregularly.

This result is consistent with known differences between vision and audition. Vision analyzes a spatial array in

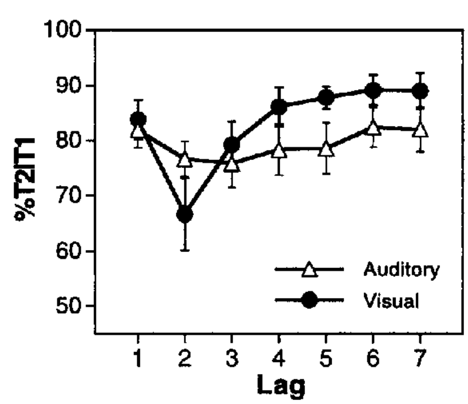

Figure 4. Correct report of the second target (T2) given correct report of the first target (T1) as a function of lag, for visual stimuli (stimulus onset asynchrony [SOA] = 135 ms) and auditory stimuli (SOA = 135 ms) in Experiment 2. Vertical bars show standard error of the mean.

parallel but handles temporal events relatively discretely, with rapid uptake and rapid shifts over time, and little or no visual persistence once a masking visual event replaces the previous stimulus. In contrast, auditory stimuli, especially speech sounds, require processing over an extended time, and the echoic buffer persists for about 2 s (Crowder, 1993). Thus, there is much greater elasticity in the timing of the later stages of processing in the case of auditory stimuli than there is for visual stimuli.

If Stage 1 processing of stimuli results in a more durable representation in the auditory case than in the visual case, that would account for the lack of a lag effect for auditory stimuli. But why should there be any auditory T2 problem at all in that case? There appears to be a general problem in encoding a second target that results in a small deficit, possibly at a stage common to visual and auditory stimuli. Our main interest was not in possible overall difficulties in encoding two targets rather than one, but rather in the temporal pattern of attention following the first of two targets. Therefore, in the remaining experiments we omitted one-target control conditions and focused on the effects of the lag between the two targets, in each condition.

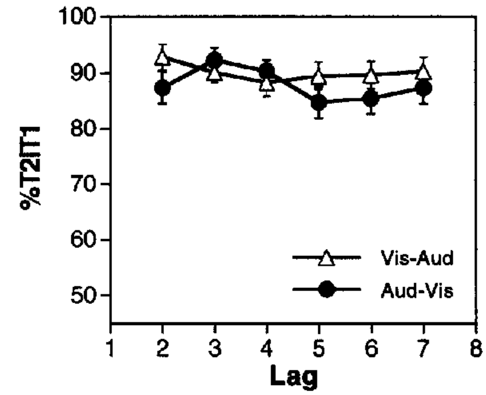

Figure 5. Correct report of the second target (T2) given correct report of the first target (T1) as a function of lag, for the two crossmodal conditions in Experiment 3; the auditory (Aud) stimuli had stimulus onset ansynchronies of 135 ms, the visual (Vis) stimuli, 120 ms. Vertical bars show standard error of the mean.

Experiment 3

In Experiment 3 we asked whether there would be a cross-modal AB, and if so, whether it would depend on the order of the two stimuli. We used the visual and auditory stimuli of Experiment 1 matched for difficulty: 135-ms auditory stimuli that included a 15-ms silent interval and 120-ms visual stimuli with no ISI.

Method

Participants. Twelve individuals from the same pool as that of Experiment 1 participated. None had been in previous AB experiments.

Design and procedure. The procedure was like that of the earlier experiments except for two changes: The modality of the sequence changed once, always on the second item after the first target (so that T1 was always followed by one distractor in the same modality), and for that reason Lag 1 targets were omitted. As before, the modality sequence was blocked, with half of the participants starting with a block in which a visual T1 and visual digit distractors changed to an auditory T2 and auditory digit distractors, and with half starting with an auditory T1 followed by a visual T2. Before each sequence a fixation cross was presented on the screen (as in the earlier experiments); when the sequence began with auditory stimuli, the fixation mark remained on the screen until the visual stimuli began.

Results and Discussion

T1 was reported correctly on 93% of the auditory trials and 92% of the visual trials ($F < 1.0$). The results for T2 are shown in Figure 5. (Note that no Lag 1 trials were included in this experiment.) As shown in the figure, performance was very good for T2, and the lag functions were essentially flat. The only significant effect was an interaction between lag and modality, $F(5, 55) = 2.40, p < .05$, apparently due to the relatively good performance at Lags 3 and 4 in the visual condition relative to the auditory condition. It is clear, however, that there was no AB in either modality. The high level of performance is unlikely to reflect a ceiling effect that concealed a true AB effect, because the same T2 visual stimulus gave a substantial AB effect in the all-visual condition in Experiment 1, and the same T2 auditory stimulus gave an overall T2 deficit (not lag dependent) in Experiment 1.

If the problem with visual stimuli is that they do not persist and are therefore often lost when their second-stage processing is delayed, one might have expected a blink when the second target was visual, even if the first one was auditory (as in one condition in Experiment 3). This assumes that an auditory target (T1) requires the same second-stage serial encoding as a visual target, setting the stage for a deficit in encoding a visual T2. But in fact there was no systematic difference between the visual-first and the auditory-first conditions—except for Lag 2, which was indeed more difficult for a visual T2. But the Lag 2 stimulus was the first visual stimulus, and perhaps did not receive the viewer's full attention.

One possible explanation for the general failure to find a blink even when the visual target came second is that participants quickly learned to concentrate first on the visual target, delaying second-stage encoding of the auditory target whether the visual target came first or second. Because the modality order was blocked, participants could have adjusted their attention relatively easily. A visual T2 target would, in this case, be encoded about as easily as a visual T1 target.

A different possibility is that second-stage encoding is carried out separately for visual and auditory stimuli, increasing total processing and memory capacity, as in a study by Scarborough (1972). In that study participants listened to a list of auditory digits slightly longer than their digit span, viewed a brief visual display of letters, and then were cued to recall either the digits or the letters. (In control conditions, the two tasks were presented separately.) The two tasks showed little mutual interference, suggesting that auditory and visual representations were maintained separately. In a cross-modal target search task, this would mean that each of the two targets would be T1 in that modality, and overall performance would be expected to be very good, as we observed in Experiment 3. This proposal is similar to one made by Shulman (1994; see footnote 1).

A third possibility is that the onset of the cross-modal stimulus stream may have captured attention, enhancing the detectability of second targets. We tested and disconfirmed this hypothesis in Experiment 5 using concurrent streams of visual and auditory stimuli that did not contain such onset cues.

In summary, the results of the three experiments indicate that the AB is modality specific: Visual sequences produced a marked and characteristic T2 lag effect, but equivalent auditory sequences produced no systematic lag effect, nor did sequences in which T1 and T2 were in different modalities.

Experiment 4

We showed that AB does not occur for auditory targets, whether detection performance for auditory and visual stimuli is matched for difficulty (Experiment 1) or for rate of presentation (Experiment 2); nor did we find any evidence for AB with cross-modal targets (Experiment 3). However, Arnell and Jolicoeur (1995) reached a different conclusion from a set of experiments that appeared to be similar to Experiments 1–3. They found significant AB effects for auditory targets and for mixed-modality targets, as well as for visual targets.[2]

Our task differed from that of Arnell and Jolicoeur (1995; in press, Experiment 2) in several respects. We used SOAs of 120 or 135 ms per item compared to their 93.3 ms per item presentation rate. As performance is likely to be dependent on presentation rate, this was a possible candidate for the difference in results we had obtained. In one of their experiments Arnell and Jolicoeur (in press, Experiment 5) varied the presentation rate of all-visual and all-auditory sequences, and found AB effects in the visual condition that increased in magnitude as SOA decreased from 150 ms to 135 ms, 120 ms, and 105 ms. With auditory sequences, there was no AB effect with SOAs of 150 or 135 ms, but there was significant and equivalent AB with SOAs of 120 ms and 105 ms. In an unpublished study we shortened the duration of the auditory stimuli of Experiments 1–3 to 120 ms, and again we failed to obtain a lag-dependent deficit for auditory–auditory targets using our procedure (Potter, Chun, & Banks, 1995). Higher presentation rates could not be tested with our stimuli because of low general comprehensibility for the auditory stimuli. Recall that auditory AB was not obtained even when overall discriminability and performance were matched for the two modalities (Experiment 1). If AB reflects limitations on central attentional mechanisms, then a similar interference effect should have been found for both visual and auditory modalities. Any dissociation between the two suggests a modality-specific locus of interference.

A major difference between our procedure and that of Arnell and Jolicoeur (1995, in press) was the nature of the task. In their key experiment (1995; Experiment 2 in Arnell & Jolicoeur, in press), they presented their participants with concurrent streams of visual and auditory letters, plus one digit in each modality that appeared at the same point in the two streams. The target (T1) was the digit in the relevant modality; the participant ignored the simultaneous digit in the other modality. The stimuli in the two streams were selected randomly and independently, so they were usually different from each other, but the onsets of the visual and auditory stimuli were synchronized. The T2 task was to decide whether an *X* was present or absent, among the letters following the T1 digit. On half of the trials there was an *X* in the designated modality. There were two within-modality conditions, visual T1–visual T2 (VV) and auditory T1–auditory T2 (AA), and two cross-modal conditions, visual T1–auditory T2 (VA) and auditory T1–visual T2 (AV). These four conditions were presented to different groups of participants. Participants attended to the stream of stimuli corresponding to the modality of the targets they were to report. For example, in the VA condition, observers monitored the visual stream for a digit T1 and then attended to the auditory stream to detect an auditory *X* (T2). All of the stimuli were presented at rates of 10.7 items per second (93.3-ms SOA).

We hypothesized that the Arnell and Jolicoeur (1995, in press) procedure entailed a capacity-demanding task switch (independent of a possible modality switch) between T1 and T2. Participants in Arnell and Jolicoeur's experiments just described (1995, in press, Experiment 2) reported the identity of the T1 digit (which was always one of four digits, 1, 2, 3, or 4) and then pressed one of two keys to indicate whether they had detected an *X* among subsequent letters. Because the letters were to-be-ignored distractors until the T1 digit was noted, a switch in perceptual set was required to detect the *X*. Allport, Styles, and Hsieh (1994) and Monsell (1996; Rogers & Monsell, 1995; also Meiran, 1996) have recently shown that a task shift, even if it is fully predictable and is cued by the nature of the stimulus, causes a deficit in performance on the first trial of the new task, relative to subsequent trials. In most of these studies, the successive trials (corresponding to T1 and T2 in the AB task) were separated by several seconds, and the main measure of performance was reaction time rather than error rate.

In two experiments, however, Allport et al. (1994, Experiments 6 and 7) presented two targets in the same trial for later report, as in AB studies. The difference, however, was that the SOA of T1 and T2 was long enough (660 ms or more, corresponding to three or more intervening items) to prevent AB (and there was no AB when the task did not switch). The variable of interest was whether the task switched between T1 and T2: A task switch was signalled by

[2] Shortly before the final revision of the present article we received a copy of Arnell and Jolicoeur (in press) from Karen Arnell; their Experiment 2 was the experiment they had reported in 1995, with minor changes. They also reported the results of several other experiments, none of which affects the conclusions of the present article except as noted.

a shift in the spatial positions of the items on the screen. (Appropriate controls showed that the shift alone did not interfere with target detection.) At that point participants had to change from, for example, searching for animal names among nonanimal names, to searching for nonanimals among animals (or searching for objects or animals smaller than a soccer ball). The main finding was that if the target was one of the first three items after the switch (and particularly if it was the first item), there was a marked reduction in the probability that the target would be reported. Thus, in a target search task similar to the AB paradigm, a task switch can produce a significant deficit that is maximal immediately after the switch and recovers after about five items, and is clearly different from AB.

In Experiment 4 we examined the possibility that task switching (TS), which presumably engages central capacity limitations common to both visual and auditory modalities (Monsell, 1996), produces a deficit that is distinct from the visual AB. To test this hypothesis we replicated the Arnell and Jolicoeur (1995, in press, Experiment 2) procedure, using the same auditory stimuli and rate of presentation that had failed to produce AB in Experiment 3 when T1 and T2 were both letters. Participants were asked to identify a lone digit target, and then to detect the presence or absence of a subsequent *X* (T2). This task imposes a category switch in perceptual set (task switch) used to select targets from distractors. The TS hypothesis predicts that a deficit like that of Arnell and Jolicoeur should be found with this procedure, regardless of modality.

Method

Participants. Thirty-six individuals from the same pool as that of Experiment 1 participated. None had been in previous AB experiments. Thirteen additional participants were replaced because they did not meet a performance criterion (see *Scoring* section).

Design and procedure. The procedure was like that of the earlier experiments except as noted. T1 was always a digit, randomly chosen from a set of digits: 2, 3, 4, 5, 6, and 9. Following the procedure of Arnell and Jolicoeur (1995), the T2 task was to determine whether an *X* had occurred or not. The lag between T1 and T2 was varied as before. The rate of presentation was 120 ms per item for both digits and letters. As in the Arnell and Jolicoeur (1995, in press, Experiment 1) experiments, two concurrent streams of stimuli were presented, one auditory and one visual. All stimuli were letter distractors except for one digit (T1) in the designated modality, and (on a random half of the trials) one *X* (T2) in the designated modality.

Each participant was tested in one of four conditions: AA, AV, VA, and VV. Participants were told what the modality or modalities of the two targets would be, and were instructed also to look at or listen to the stimuli in the irrelevant modality. Participants reported the identity of the digit by typing the corresponding arabic numeral on the keyboard and then pressed the *F* key (covered with a sticker labeled *N* for *no*) or the *J* key (with a label *Y* for *yes*) to indicate whether an *X* had appeared; they were also instructed that they could press the space bar if they were uncertain of whether an *X* had appeared, and such responses were counted as "no" responses.

Scoring. Two criteria were set for inclusion of a participant in the data analyses: T1 had to be reported correctly on at least 60% of the trials, and the false alarm (FA) rate on T2 (the percentage of false yes responses when no *X* had been presented) had to be no higher than 30% (the overall mean FA rate among included participants was 15.8%). The reason for adopting an FA criterion was to avoid including participants with an excessively low criterion for making a detection response, which would have indicated that the correct detections were contaminated by guessing. A total of 13 participants (27%) were replaced because they did not meet one of the criteria: 4 in the AA group (4 who did not meet the FA criterion, 1 of whom also did not meet the T1 criterion); 5 in AV (3 T1s, 1 FA, and 1 both); 2 in VA (both FAs); and 2 in VV (both FAs). Overall, 7 did not meet the FA criterion, 3 the T1 criterion, and 3 both criteria.

Results and Discussion

Report of T1. T1 was reported correctly on 83.9%, 75.6%, 94.8%, and 90.1% of the trials in the AA, AV, VA, and VV conditions, respectively. Whether an *X* was present on a given trial had no effect on T1 report ($F < 1$), and the interaction with conditions was not significant, $F(3, 32) = 1.37$. A 2 (T1 modality) × 2 (T2 modality) × 7 (lag) ANOVA was carried out on T1 performance. Participants were more accurate reporting visual T1s than auditory T1s (92.5% vs. 79.8%, respectively), $F(1, 32) = 24.87$, $p < .001$. T1 performance was higher when followed by an auditory probe than by a visual probe (89.4% vs. 82.9%, respectively), $F(1, 32) = 6.50$, $p < .05$. No other main effects or interactions approached significance.

Report of T2|T1. As shown in Figure 6, the VV conditions once again showed a marked AB effect, that is, a nonmonotonic lag effect with the lowest performance at Lag 2. However, in striking contrast to our earlier findings, a lag effect was now obtained in the other three conditions as well. Consistent with the TS hypothesis, this deficit was monotonic, with the lowest performance at Lag 1. An ANOVA was carried out on the percentage of correct detections of the *X*, on trials in which an *X* was present and the digit had been reported correctly. In an analysis of T1 Modality × T2 Modality × Lag, the main effect of T1 modality on T2 detection was not significant; T2 accuracy was 82.1% when T1 was visual and 75.3% when T1 was auditory ($p > .10$). Overall, visual T2s were more difficult to detect than auditory T2s (74.6% vs. 82.8%, respectively), $F(1, 32) = 4.11$, $p = .051$. The interaction between T1 modality and T2 modality was not significant, $F(1, 32) = 2.73$, $p > .10$. There was a main effect of lag, $F(6, 192) = 19.28$, $p < .001$, and significant two-way interactions between lag and T1 modality, $F(6, 192) = 3.26$, $p < .01$, and lag and T2 modality, $F(6, 192) = 6.10$, $p < .01$, as well as a significant three-way interaction among lag, T1 modality, and T2 modality, $F(6, 192) = 5.93$, $p < .001$.

In separate analyses of the four modality groups, lag was significant for the AA, AV, and VV conditions: for AA, $F(6, 48) = 4.94$, $p = .001$; for AV, $F(6, 48) = 2.31$, $p < .05$; and for VV, $F(6, 48) = 20.8$, $p < .001$. Lag was marginally significant for the VA condition, $F(6, 48) = 2.11$, $p < .07$. The mean false-yes rates on probe-absent trials, given that T1 had been reported correctly, were 17.6%, 8.9%, 14.1%, and 15.4% in the AA, AV, VA, and VV conditions, respectively, $F(18, 192) = 2.50$, $p < .08$.

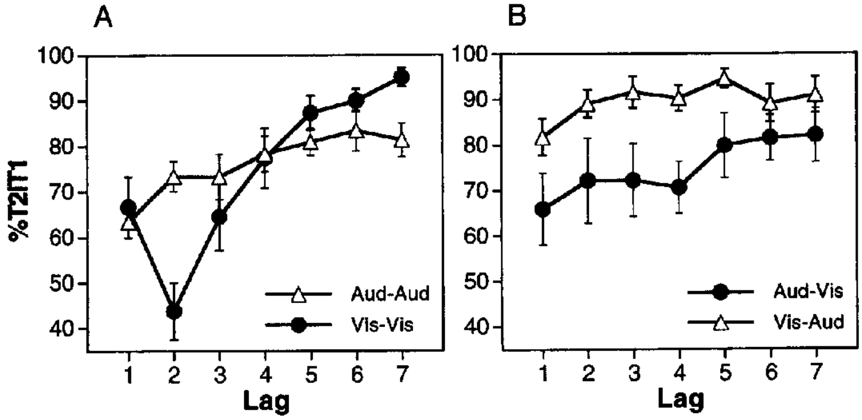

Figure 6. Correct report of the second target (T2) given correct report of the first target (T1) as a function of lag in Experiment 4, with two simultaneous sequences of visual (Vis) and auditory (Aud) items (stimulus onset asynchrony = 120 ms) and instructions to identify T1 and detect *X* (T2) in specified modalities, for (A) the groups with two Vis or two Aud targets, and (B) the groups with one Vis and one Aud target. Only trials on which an *X* was presented are included. Vertical bars show standard error of the mean.

Thus, using the T1 and T2 tasks of Arnell and Jolicoeur (1995, in press, Experiment 2) with dual streams of stimuli, but with the same materials used in our earlier experiments, we replicated the auditory and cross-modal deficits they had reported. Before discussing this result further, we report another experiment designed to control for another variable: the use of dual streams of stimuli in Experiment 4 but not in the earlier experiments.

Experiment 5

Introducing a task switch in Experiment 4 produced cross-modal and all-auditory T2 deficits that were not present in Experiments 1–3. However, there were several differences in procedure that complicate a comparison of the results across the experiments. First, single streams of stimuli were used in Experiments 1–3, whereas the presentation of dual streams (one for each modality) may have increased the overall load of the task in Experiment 4. Second, the sudden onset of the opposite modality stream after T1 in Experiment 3 may have captured attention, facilitating T2 identification (e.g., Posner, Nissen, & Klein, 1976).[3]

We conducted Experiment 5 to eliminate these confounds. All four conditions (AA, AV, VA, and VV) of Experiment 4 were replicated using dual streams of stimuli. The only critical difference was that the task-switch aspect of the task was eliminated by having participants search for two letter targets among digits, instead of switching perceptual sets between a digit T1 and a letter T2. Our prediction was that there would be the usual AB effect when both targets were visual, but that there would be no AB effect in the other three conditions, replicating the results from Experiments 1–3. Such a pattern would reinforce the conclusion that the T2 deficits in the auditory and cross-modal conditions in Experiment 4 (and in Arnell & Jolicoeur, 1995, in press, Experiment 2) were due to a task switch, rather than to the use of dual streams versus single streams of stimuli.

Method

Participants. There were 36 participants from the same pool as that of Experiment 1. None had been in previous AB experiments.

Design and procedure. The design and procedure were identical to those of Experiment 4 except as follows. T1 and T2 were both nonidentical letters, and there were no trials without a T2. The letter targets and digit distractors were drawn randomly from 18 letters and six digits, respectively, as in Experiments 1–3. Replicating Experiment 4, two concurrent streams of stimuli were presented, one auditory and one visual. All stimuli were presented at SOAs of 120 ms, as in Experiment 4. Participants were told what the modality or modalities of the two letter targets would be, and were instructed also to look at or listen to the stimuli in the irrelevant modality; they reported the identities of the two target letters by typing them, in the order in which they occurred, after the sequence had been presented, or pressed the space bar if they had no idea about the identity of the target.

Results and Discussion

Targets were scored as correct regardless of order of report, as in Experiments 1–3. T1 was reported correctly on 82.2%, 81.4%, 91.9%, and 82.8% of the trials in the AA, AV, VA, and VV conditions, respectively, $F(3, 32) = 2.84$, $p < .06$. As in Experiment 4, visual T1s ($M = 87.4\%$) were easier to report than auditory T1s ($M = 81.8\%$); the effect of T1 modality approached significance, $F(1, 32) = 3.65$, $p < .07$. There was no effect of T2 modality on T1 performance ($p > .10$), although there was a trend for visual T2s to interfere more with T1 performance ($M = 82.1\%$) than

[3] We thank Tram Neill and Kimron Shapiro for pointing out these factors and suggesting Experiment 5.

auditory T2s ($M = 87.0\%$), replicating the pattern obtained in Experiment 4.

As shown in Figure 7A, there was a marked AB effect in the VV condition, with a substantial benefit at Lag 1 and a maximal deficit at Lag 2. As predicted, no AB effect was found in the AA and AV conditions. There was a significant AB effect in the VA condition, but it was small in magnitude. In an analysis with T1 Modality × T2 Modality × Lag as variables, T1 modality had no main effect on T2 performance; T2 accuracy was 86.5% with visual T1s and 85.0% with auditory T1s ($F < 1$). Likewise, visual and auditory T2s were similar overall (84.9% vs. 86.6%; $F < 1$). There was a significant interaction between T1 modality and T2 modality, $F(1, 32) = 7.58, p < .01$. There was a main effect of lag, $F(6, 192) = 8.04, p < .001$, and significant two-way interactions between lag and T1 modality, $F(6, 192) = 4.05$, $p < .001$, and lag and T2 modality, $F(6, 192) = 2.38$, $p < .05$, as well as a significant three-way interaction, $F(6, 192) = 4.82, p < .001$.

In separate analyses of the four modality groups, lag was significant for the VV group, $F(6, 48) = 9.88, p < .001$, and the VA group, $F(6, 48) = 3.26$, $p < .01$, but was not significant for the AA ($F < 1.03$) and AV ($F < 1$) groups.

The results of Experiment 5 confirm the hypothesis that the auditory and cross-modal deficits reported by Arnell and Jolicoeur (1995, in press) and replicated in Experiment 4 can be accounted for by the requirement to switch selective sets between T1 and T2. When participants searched for two targets of the same type in Experiment 5, little or no AB was found for auditory or cross-modal stimuli. But when T1 and T2 are defined differently and require a switch in perceptual set (as in Experiment 4 and in the Arnell & Jolicoeur experiments [1995, in press, Experiment 2]), there is a transient deficit whatever the modality or modalities of the two targets.

These results suggest that there are two distinct sources of interference between the processing of one target and a subsequent one. One source of interference is strictly visual, producing a strong AB effect on the second of two visual targets presented in RSVP. The time course of this interference is U-shaped, with a lack of impairment for T2 when it immediately follows T1 (Lag 1). (As discussed earlier, the first target and the immediately following item may be processed together in Stage 2.) This bottleneck we classify as a standard AB. In contrast, it is striking that the T2 deficit observed in Experiment 4 was maximal at Lag 1—except in the VV condition. This result suggests that no advantage can be taken of temporal proximity when the two targets require different kinds of processing except when the two targets are both visual, and even then the advantage is muted.

The second source of interference is amodal, and the present experiments suggest that it can be triggered by a distinct process, TS. Thus we concur with suggestions by Arnell and Jolicoeur (1995, in press) that an amodal, central bottleneck may play a role in the processing of rapidly presented stimuli. However, the present experiments strongly suggest that this amodal deficit is distinct from the bottleneck that is invoked in the standard attentional blink with visual targets. Note that the VV condition in Experiment 4 and in Arnell and Jolicoeur's experiment (in press, Experiment 2) should reflect components of both standard visual AB and the TS effect. The results of Experiment 4 support that prediction: The VV condition showed a larger lag-dependent deficit than the other three conditions, and the pattern appears to be a combination of Lag 1 sparing (the standard AB pattern) and a Lag 1 deficit (which tends to be maximal at Lag 1 with TS). In Arnell and Jolicoeur's experiment (in press, Experiment 2) the VV condition likewise produced a large deficit, maximal at Lags 1 and 2 (which were equal—a hint of Lag 1 sparing).

Thus, in Experiment 5, simply changing the task to one in which two letters were to be reported eliminated the significant lag effect observed in the AA and AV conditions in Experiment 4. The critical factor in producing a deficit

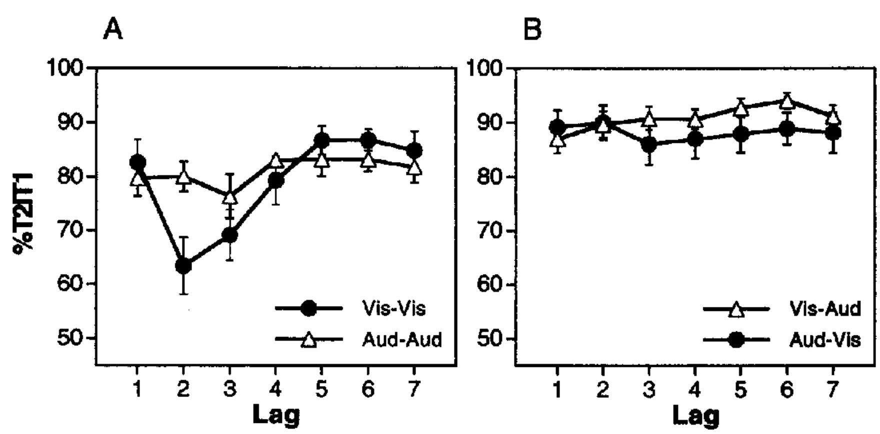

Figure 7. Correct report of the second target (T2), given correct report of the first target (T1) as a function of lag in Experiment 5, with two simultaneous sequences of visual (Vis) and auditory (Aud) items (stimulus onset asynchrony = 120 ms), when the task was to report two letters among digit distractors, for (A) the groups with two Vis or two Aud targets, and (B) the groups with one Vis and one Aud target. Vertical bars show standard error of the mean.

after an auditory target thus appears to be a task switch or switch in set between T1 and T2. Although the significant deficit in the VA condition suggests that a visual T1 is capable of producing a cross-modal deficit, inspection of Figure 7B shows that the deficit is relatively small and that the greatest deficit is at Lag 1, unlike the VV deficit.

Experiment 6

A difference between Experiment 4 and Experiments 1–3 and 5, in addition to the T2 task, was that in Experiment 4 T1 was a digit among letters, rather than a letter among digits. The digit stimuli we used were somewhat more difficult to identify than the letters, when the target was auditory. To ensure that this factor was not accounting for the difference between the results of Experiments 4 and 5, in Experiment 6 we replicated the AA condition of Experiment 4, but replaced the *X* (T2) with a second digit to be identified and reported. (In place of the *X*-absent trials in Experiment 4, a second digit target was added in Experiment 6, with lag counterbalanced.) The participants were instructed to report the two auditory digits among auditory letters, while also watching a visual stream of letters. We predicted that, just as in the AA conditions of Experiments 1, 2, and 5, there would be no evidence for a lag-dependent T2 deficit.

Method

Participants. There were 16 participants from the same pool as that of Experiment 1. None had been in previous AB experiments. Two additional participants were replaced because they did not meet the criterion for at least 60% correct report of T1.

Design and procedure. The procedure was like that of Experiment 4's AA condition except that T2 was also a digit, like T1, and there were no trials without a T2. A simultaneous sequence of visual letter distractors was presented, as in Experiment 4. There were 336 trials, 48 at each of the seven lags (1–7). T1 and T2 were selected randomly without replacement from the same set of six digits used in Experiment 4. In other respects the method was like that of previous experiments.

Results and Discussion

Target report performance was scored regardless of order of report (an analysis using the stricter criterion that the report had to be in the correct order gave similar results). T1 was reported correctly on 75.5% of the trials. Figure 8 shows the percentage of T2 reports at each lag, given correct report of T1. An ANOVA showed no significant effect of lag, $F(6, 90) < 1.0$.

Thus, the results of Experiment 6 reinforce the conclusion from Experiments 4 and 5 that the critical factor in producing a lag-dependent auditory or cross-modal deficit is a task switch or switch in set between T1 and T2. Without this switch, lag-dependent deficits are robust only in all-visual conditions.

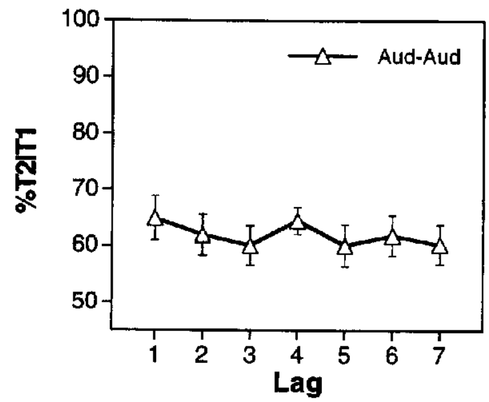

Figure 8. Correct report of the second target (T2), given correct report of the first target (T1) as a function of lag in Experiment 6, with two simultaneous sequences of visual and auditory (Aud) items (stimulus onset asynchrony = 120 ms) and instructions to identify two Aud digits among letter distractors. Vertical bars show standard error of the mean.

General Discussion

In the present experiments participants searched for two targets presented among a rapid serial sequence of distractors. The targets were both visual, both auditory, or one was auditory and the other was visual: In all cases, the modality of the targets was blocked or between subjects. As shown in previous studies, a robust AB effect was found for visual targets. Correct report of T1 was accompanied by a marked impairment in report of a second target appearing within 200–500 ms when the task was to report two letters presented among digit distractors. No such AB effect was found for two targets that were presented auditorily (Experiments 1, 2, 5, and 6), nor did presentation of a target in one modality impair performance for a target appearing in a different modality (Experiments 3 and 5). Although there was sometimes a general impairment for reporting two targets as opposed to one (Duncan, 1980), for auditory and cross-modal stimuli this interference was not dependent on lag, in striking contrast to the strong lag effects obtained for visual targets. This pattern of results was replicated whether visual and auditory target detection was matched for difficulty (Experiments 1 and 3) or for rate of presentation (Experiments 2 and 5).

A different pattern of results was obtained in Experiment 4, which replicated a study by Arnell and Jolicoeur (1995, in press, Experiment 2). A significant interference effect for auditory and cross-modal as well as for visual stimuli was obtained when a switch in perceptual set or task was introduced between the two targets. In this experiment participants were instructed to search for a digit and then a probe letter (target category switch): In all conditions a lag-dependent deficit was observed when the second target appeared within 500 ms. In Experiment 5, which used the same method as Experiment 4 but with two letters as the targets, a large effect was obtained only in the VV condition, and no lag effects were observed in the AA and AV conditions. (A slight lag effect was observed in the VA condition.) Experiment 6 replicated the lack of a lag effect for AA targets in Experiment 5, using digits as targets as in Experiment 4's T1 task.

These results support the hypothesis that there are two distinct attentional deficits, one of which is specific to visual processing, as evidenced by the absence or near-absence of such an effect for auditory or cross-modal targets when there is no task switch. This visual deficit we term *standard AB.* The two-stage model of AB (Chun & Potter, 1995b) proposes that stimuli appearing at rates of up to 10 per second can be fully identified (Stage 1 processing), but must pass through a limited-capacity stage of processing for further analysis and consolidation of the target for subsequent report (Stage 2). Stimuli appearing while Stage 2 is engaged by a previous target must wait. A key assumption is that visual items waiting in the Stage 1 (postcategorical) buffer are unstable and readily interfered with by subsequent items, so that they may be lost by the time Stage 2 is freed up. A recent study by Giesbrecht and Di Lollo (in press) with visual stimuli has shown that items appearing during the blink interval or in conditions of inattention are highly susceptible to erasure by a subsequent item.

There are two possibly related reasons that standard AB is not observed for auditory or cross-modal targets: Auditory targets do not require a Stage 2 process, or the echoic buffer is less susceptible to erasure from subsequent items than is the visual Stage 1 buffer, or both. There is independent evidence for the robustness and relatively long duration of the echoic buffer (Crowder, 1993), supporting the second of the two explanations for the absence of a standard AB deficit with auditory and cross-modal targets. In the AV case, participants may opt to process the visual target (T2) first, while the auditory (T1) target is held in echoic memory. (Arnell & Jolicoeur, in press, reported that some of their participants adopted that strategy in an experiment in which the T1 auditory signal was a distinctive high tone and T2 was a visual probe.)

A different type of interference is triggered when a switch in task or perceptual set is required from T1 to T2. This TS deficit is amodal, suggesting a locus or process that is common to the visual and auditory modalities. In Experiment 4, in which participants were required to switch from digit identification to letter detection, there was interference with T2 in all four modality conditions. There were two facets to the task switch, one the switch in selective set from digits to a specified letter (*X*), and the other the switch from ignoring letter distractors to attending to the letter *X* (once the T1 digit target was identified). TS, we hypothesize, is associated with a temporary impairment in target detection. This hypothesis is consistent with the evidence for TS costs between trials (Allport et al., 1994; Meiran, 1996; Rogers & Monsell, 1995) and is directly supported by within-trial TS deficits shown in Allport et al.'s study, described earlier. Meiran (1996) suggested that a change in task demands requires the recruitment of executive control functions that reconfigure the system for the new task. Likewise, the TS deficit reported in our Experiment 4 may reflect limitations in reconfiguring the perceptual set from digits to the letter *X*. (Note that it is an open question whether the TS costs result from a serial bottleneck or from competition among parallel processes for limited resources; either could produce the TS effects observed here.)

We noted earlier, in justifying the omission of T1 as a control condition in Experiment 1, that it was difficult for participants in a pilot experiment (Chun, 1992) to report only the second of two letter targets: They showed almost as large an AB deficit as they did when reporting both targets. In contrast, in experiments that have used the *X* detection task (e.g., Raymond et al., 1992, and Arnell & Jolicoeur, 1995, in press), instructing the participant to ignore T1 and report only T2 (or, in this example, its presence or absence) eliminates the AB effect. This difference in the ability to focus selectively on T2 is a further indication that a task difference, and hence a task switch, is involved in one case and not the other.

While this article was undergoing revision, a new study was published that also reported a lack of AB for cross-modal targets (Duncan, Martens, & Ward, 1997). However, in contrast to the present study, they did find an attentional deficit for two auditory targets. There were several differences between their task and the present dual-stream tasks. In their task participants detected two monosyllabic words; although no categorical task switch was involved, the pair of words to be discriminated in a given channel was different from the pair in the other channel. The SOA was 250 ms rather than 120–135 ms. Perhaps the most critical difference in procedure was that they used two concurrent, auditory streams of stimulation, one in a low voice and one in a high voice, whereas we used a single stream of stimuli in each modality. Their report of a deficit in identifying stimuli from different auditory channels is consistent with classic findings showing that people can only attend to a single auditory channel (e.g., Treisman & Davies, 1973). A comparison of their study and ours outlines critical boundary conditions of auditory, visual, and cross-modal AB. Visual AB occurs for single streams of input (as well as for dual streams), whereas auditory AB appears to require dual streams of auditory stimulation for within-modality capacity limitations to be revealed. In either case, limitations in one modality do not interfere with processing in the other unless a task switch is introduced.

Thus, both AB and TS deficits are found in serial target search tasks. The first difference between the two deficits is that standard AB is found only when both targets are visual, whereas the TS deficit is found whatever the modalities of the two targets. (Note that the studies reported here only investigated auditory and visual modalities, so we cannot be sure what would happen with other modalities.)

The second distinctive marker is performance at Lag 1. In the visual modality, first-stage identification of the first target evokes an attentional process that is temporally imprecise, so that in Stage 2 the first target is frequently accompanied by the immediately following visual stimulus, be it a random mask, a distractor, or a second target. In the latter case both targets are processed together, so that the normal AB effect is reduced or eliminated: We term this effect *Lag 1 sparing.* Thus, standard visual AB exhibits a U-shaped function of lag, with the poorest performance at Lag 2 (or 3) rather than Lag 1. In contrast, interference produced by TS is monotonic: maximal at Lag 1 and recovering with increasing lag. This TS pattern of performance was shown not only in Experiment 4, but also in the experiments reported by Arnell and Jolicoeur (1995, in press). Another striking example of T2 monotonicity was found in a visual study by Joseph, Chun, and Nakayama (1997). In their experiments T1 was a white letter appearing within an RSVP stream of black letters. Their T2 task involved determining the presence or absence of an oddball target orientation (Gabor stimuli were used). The demands of switching the selection set from white letter identification to orientation oddball detection resulted in a maximal impairment at Lag 1 in that task. Note that we are not proposing that performance at Lag 1 is the main criterion for distinguishing between AB and the task-switch deficit. However, our working hypothesis is that visual AB in RSVP tasks will involve some measure of Lag 1 sparing, which may be overridden by TS. Further research is needed to fully specify the conditions that produce Lag 1 sparing versus monotonicity.

The third and most important distinction is in the proposed locus of the two types of interference. By now there is a rich body of evidence that indicates that visual AB is occurring postcategorically. This is consistent with the claim of the two-stage model (Chun & Potter, 1995b) that Stage 1 representations are conceptual. There is substantial empirical support, both behavioral and neurophysiological, for the assumption that there is semantic activation of items that are attentionally blinked (Luck et al., 1996; Maki et al., 1997; Shapiro et al., 1997). Thus, visual AB appears to be occurring somewhere between identification and encoding into short-term memory, consistent with Chun and Potter's (1995b) proposal and recently Jolicoeur's (in press) proposal that Stage 2 represents a consolidation process required for overt report.

Other prominent models of AB include the visual short-term memory (VSTM) interference model of Shapiro and Raymond (1994) and the central interference theory of Jolicoeur and Dell'Acqua (1996), which we consider in turn. Shapiro and Raymond attributed visual AB to retrieval interference between T1 and T2 in VSTM. The present data do not allow for a distinction between the two-stage model and the VSTM model, as both models propose that AB is specific to the visual modality. The two-stage model and the VSTM model make similar predictions for AB in a variety of other tasks, and recent efforts have been made to promote further theoretical convergence (Shapiro, Arnell, & Raymond, 1997). For example, the VSTM model or "unified account" now incorporates the two-stage model's assumption that AB occurs because conceptual representations of T2 are lost prior to consolidation, as a consequence of interference from stimuli subsequent to T2. This was based on recent evidence showing that attention is needed to prevent items from being replaced by subsequent stimuli (Enns & Di Lollo, 1997; Giesbrecht & Di Lollo, in press; see also Chun, 1997a) and it is consistent with Potter's (1993) evidence for very short-term conceptual memory.

However, differences between the two models still exist. A detailed discussion of this evidence goes beyond the scope of this article, but a brief review here will help constrain future interpretations of the AB deficit. First, the two-stage model posits that T1 processing difficulty (within the visual modality and restricted to perceptual processing as influenced by factors such as masking) determines the magnitude of AB, whereas the VSTM model explictly claims that T1 processing difficulty does not correlate with blink magnitude (Shapiro et al., 1994). This claim was based on a correlation analysis of four experiments, one of which included a target condition that was not strictly visual (gap duration detection). The VSTM model's assumption has been disconfirmed in a number of studies that showed that T1 processing difficulty is an important factor in AB (Grandison, Ghirardelli, & Egeth, 1997; Jolicoeur, in press; Seiffert & Di Lollo, 1997). In particular, Seiffert and Di Lollo presented a meta-analysis of 27 published experiments showing that T1 identification accuracy correlates negatively with blink magnitude. Grandison et al. (1997) argued that T1 processing difficulty is probably the only general parameter that characterizes triggering conditions of AB. Furthermore, Jolicoeur (in press) has recently shown that T2 performance not only correlates negatively with T1 accuracy, but also negatively covaries with the time it takes to respond to T1. Second, the VSTM model proposes that increased T1–T2 similarity increases the magnitude of the effect, linking AB to a related effect known as repetition blindness (RB). RB is a deficit in the report of a second identical (or similar) target (Kanwisher, 1987). However, AB was shown to occur without RB, and RB was demonstrated without accompanying AB (Chun, 1997b). Such a double dissociation does not support the notion that T1–T2 similarity is a critical determinant of AB.

A new account of AB has recently been put forth by Jolicoeur and Dell'Acqua (1996), who proposed that AB for T2 occurs when limited capacity resources are occupied by processing and consolidation of T1: a hypothesis very similar to that of Chun and Potter (1995b). However, Jolicoeur and Dell'Acqua additionally argued that the resources consumed during the AB interval are the same central resources that are used in a variety of other cognitive operations such as response selection (psychological refractory period, PRP). Our data suggest that Stage 2 processing recruits visual attentional resources that are distinct from those used for auditory processing or response selection (see also Duncan et al., 1997), supporting the two-stage model or the VSTM model over the more generic account put forth by Jolicoeur and Dell'Acqua. Although our data suggest that AB reflects capacity limitations within the visual modality only, we re-emphasize that this independence does not imply that visual target detection performance is immune from other types of cognitive interference. For instance, Jolicoeur (in press) has shown that imposing a speeded response requirement on T1 increases AB on T2. This result is still consistent with our view that interference in processing a second target can occur at a multitude of processing stages, such as response selection in PRP or the TS cost reported here.

The actual locus of TS interference is less clear, and further research is warranted. The lack of Lag 1 sparing and the amodal character of the TS effect suggest a link with the PRP, which is also monotonic, producing maximal interference at the shortest lags. However, Pashler (1989) showed that response selection in a primary task (which triggers PRP) did not impair accuracy in a secondary task, detection of a masked digit (corresponding to T2 here). Moreover, Pashler (1994) has also shown that the response selection bottleneck is present even when the task is held constant. Thus, for now, the literature suggests that the TS effect reported here is distinct from the response selection bottleneck that triggers PRP. The TS deficit is directly linked to costs involved in switching selective set, as shown by deficits found in such studies as those of Allport et al. (1994), Meiran (1996), and Rogers and Monsell (1995). The magnitude of the interference effect decreases monotonically as a function of intertrial interval, although it should be noted that the ISIs used in most of those studies were all considerably longer than in the present studies, inasmuch as an overt response was made after each stimulus. It is likely that in all of these cases the switch to a new perceptual set takes time, resulting in slowed responses to the next stimulus in the TS studies mentioned, and second-target misses in search tasks like those we have studied (see also Allport et al., 1994, Experiments 6 and 7, described earlier). In any case, there is a whole class of interference effects that result in central attention, memory, or response selection limitations, as suggested by Jolicoeur and Dell'Acqua (1996), Monsell (1996), and by the work of Pashler and his colleagues (Carrier & Pashler, 1995; Pashler, 1989, 1994; see Pashler, 1992, 1993, for reviews).

An important question, and a potential clue to the loci of the two deficits, is whether AB and TS deficits are additive or interactive. In Experiment 4's VV condition there was a suggestion of additivity, in that the largest deficit was obtained in that condition and Lag 1 was somewhat less impacted than Lag 2 (see Figure 6). This would suggest that the two effects occur in different stages. However, as already noted, Arnell and Jolicoeur (1995, in press) did not find any Lag 1 sparing in that condition, and in their experiments the AV condition had as marked a deficit as the VV condition. Thus, the loci of the two effects cannot yet be determined with any confidence.

The present claim is that changing from the detection and identification of one sort of target, such as a digit or shape or white letter, to the detection or identification of another sort of target, such as an *X* among other letters, results in a TS deficit that is distinct from the standard visual AB. Although a few studies have used the same task for T1 and T2 (Broadbent & Broadbent, 1987; Chun & Potter, 1995b), much of the initial work on attentional deficits, including the work that led to the term *attentional blink* (Raymond et al., 1992), involved a task switch of this kind. All of these studies, however, used all-visual tasks, so that one would expect to see both AB and TS deficits, just as in Experiment 4's all-visual condition. A reconsideration of some of the variables explored in the earlier studies might shed further light on factors that affect AB and TS effects differentially.

The demonstration of two different types of attentional deficit is consistent with the commonly accepted notion of multiple bottlenecks in the human information processing stream (Pashler, 1992). Further research should help to clarify the distinction introduced here between the AB and the TS deficit, as well as the PRP.

References

Allport, D. A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umilta & M. Moscovitch (Eds.), *Attention and performance XV* (pp. 421–452). Hillsdale, NJ: Erlbaum.

Arnell, K., & Jolicoeur, P. (1995, November). *Allocating attention across stimulus modalities: Evidence from the attentional blink phenomenon.* Poster session presented at the 36th Annual Meeting of the Psychonomic Society, Los Angeles, CA.

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

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

Carrier, L. M., & Pashler, H. (1995). Attentional limits in memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1339–1348.

Chun, M. M. (1992). [Report of only the second of two letters in an RSVP stream of digits]. Unpublished raw data.

Chun, M. M. (1997a). Temporal binding errors are redistributed by the attentional blink. *Perception & Psychophysics, 59,* 1191–1199.

Chun, M. M. (1997b). Types and tokens in visual processing: A double dissociation between the attentional blink and repetition blindness. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 738–755.

Chun, M. M., & Potter, M. C. (1995a). [Null effects of post-T2 list length in the attentional blink]. Unpublished raw data.

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

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, England: Oxford University Press.

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

Duncan, J., Martens, S., & Ward, R. (1997, June 19). Restricted attentional capacity within but not between sensory modalities. *Nature, 387,* 808–810.

Enns, J. T., & Di Lollo, V. (1997). Object substitution: A new form of masking in unattended visual locations. *Psychological Science, 8,* 135–139.

Giesbrecht, B. L., & Di Lollo, V. (in press). Beyond the attentional blink: Visual masking by object substitution. *Journal of Experimental Psychology: Human Perception and Performance.*

Grandison, T. D., Ghirardelli, T. G., & Egeth, H. E. (1997). Beyond similarity: Masking of the target is sufficient to cause the attentional blink. *Perception & Psychophysics, 59,* 266–274.

Hunt, S. M. J. (1994). MacProbe: A Macintosh-based experimenter's workstation for the cognitive sciences. *Behavior Research Methods, Instruments, & Computers, 26,* 345–351.

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

Jolicoeur, P., & Dell'Acqua, R. (1996, November). *Attentional and structural constraints on memory encoding.* Paper presented at the 37th Annual Meeting of the Psychonomic Society, Chicago, IL.

Joseph, J. S., Chun, M. M., & Nakayama, K. (1997, June 19). Attentional requirements in a preattentive feature search task. *Nature, 387,* 805–808.

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

Kanwisher, N. (1994, November). *Repetition blindness: A reduction in sensitivity for detection of repeated items.* Poster presentation at the 35th Annual Meeting of the Psychonomic Society, St. Louis, MO.

Luck, S. J., Vogel, E. K., & Shapiro, K. L. (1996, October 17). Word meanings can be accessed but not reported during the attentional blink. *Nature, 383,* 616–618.

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 and Performance, 23,* 1014–1034.

Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1423–1442.

Monsell, S. (1996). Control of mental processes. In V. Bruce (Ed.), *Unsolved mysteries of the mind: Tutorial essays in cognition* (pp. 93–148). Hove, England: Erlbaum

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. (1992). Attentional limitations in doing two tasks at the same time. *Current Directions in Psychological Science, 1,* 44–48.

Pashler, H. (1993). Doing two things at the same time. *American Scientist, 81,* 48–55.

Pashler, H. (1994). Overlapping mental operations in serial performance with preview. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 47*(A), 161–191.

Posner, M. I., Nissen, M. J., & Klein, R. (1976). Visual dominance: An information-processing account of its origins and significance. *Psychological Review, 83,* 157–171.

Potter, M. C. (1993). Very short-term conceptual memory. *Memory & Cognition, 21,* 156–161.

Potter, M. C., Chun, M. M., & Banks, B. S. (1995). [Report of auditory letters among auditory digit distractors (SOA = 120 ms)]. Unpublished raw data.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance, 18,* 849–860.

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

Scarborough, D. (1972). Memory for brief visual displays of symbols. *Cognitive Psychology, 3,* 408–429.

Seiffert, A. E., & Di Lollo, V. (1997). Low-level masking in the attentional blink. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 1061–1073.

Shapiro, K., Arnell, K., & Raymond, J. (1997). The attentional blink: A view on attention and a glimpse on consciousness. *Trends in Cognitive Science, 1,* 291–296.

Shapiro, K., Driver, J., Ward, R., & Sorensen, R. E. (1997). Priming from the attentional blink: A failure to extract visual tokens but not visual types. *Psychological Science, 8,* 95–100.

Shapiro, K. L., & Raymond, J. E. (1994). Temporal allocation of visual attention: Inhibition or interference? In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory mechanisms in attention, memory and language* (pp. 151–188). Boston, MA: Academic Press.

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 and Performance, 20,* 357–371.

Shulman, H. (1994, November). *The attention blink in rapid serial presentation and rapid auditory presentation.* Paper presented at the 35th Annual Meeting of the Psychonomic Society, St. Louis, MO.

Treisman, A. M., & Davies, A. (1973). Divided attention to ear and eye. In S. Kornblum (Ed.), *Attention and performance IV* (pp. 101–117). New York: Academic Press.

Weichselgartner, E., & Sperling, G. (1987). Dynamics of automatic and controlled visual attention. *Science, 238,* 778–780.

Received February 7, 1997
Revision received September 29, 1997
Accepted September 29, 1997 ■