VU Research Portal

Audiovisual events capture attention: Evidence from temporal order judgments

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

document version
Publisher's PDF, also known as Version of record

Link to publication in VU Research Portal

Audiovisual events capture attention: Evidence from temporal order judgments

Erik Van der Burg

Department of Cognitive Psychology,
Vrije Universiteit Amsterdam,
The Netherlands

Christian N. L. Olivers

Department of Cognitive Psychology,
Vrije Universiteit Amsterdam,
The Netherlands

Adelbert W. Bronkhorst

TNO Human Factors, Soesterberg, The Netherlands, &
Department of Cognitive Psychology,
Vrije Universiteit Amsterdam,
The Netherlands

Jan Theeuwes

Department of Cognitive Psychology,
Vrije Universiteit Amsterdam,
The Netherlands

Is an irrelevant audiovisual event able to guide attention automatically? In Experiments 1 and 2, participants were asked to make a temporal order judgment (TOJ) about which of two dots (left or right) appeared first. In Experiment 3, participants were asked to make a simultaneity judgment (SJ) instead. Such tasks have been shown to be affected by attention. Lateral to each of the dots, nine irrelevant distractors continuously changed color. Prior to the presentation of the first dot, a spatially non-informative tone was synchronized with the color change of one of these distractors, either on the same side or on the opposite side of the first dot. Even though both the tone and the distractors were completely irrelevant to the task, TOJs were affected by the synchronized distractor. TOJs were not affected when the tone was absent or synchronized with distractors on both sides. SJs were also affected by the synchronized distractor, ruling out an alternative response bias hypothesis. We conclude that audiovisual synchrony guides attention in an exogenous manner.

Keywords: temporal order judgment, simultaneity judgment, multisensory integration, attention

Introduction

Information from different senses is often integrated in a unified percept when presented simultaneously or in close succession (for reviews, see, e.g., Spence, 2007; and Welch & Warren, 1980). For example, with regard to vision and audition, identification of visual stimuli is improved by accompanying auditory stimuli (Doyle & Snowden, 2001; Olivers & Van der Burg, 2008; Vroomen & De Gelder, 2000). Also, the perceived location of a sound is shifted toward the location of a simultaneously nearby presented visual stimulus—a finding known as the ventriloquism effect (Bertelson & Radeau, 1981; Thomas, 1941).

An important but largely unexplored question is how multisensory integration affects the competition for selective attention between multiple objects. The majority of studies reporting multisensory integration have typically used single visual and auditory events at a time (Spence, 2007). Recently, Van der Burg, Olivers, Bronkhorst, and Theeuwes (in press) investigated audiovisual multisensory integration in more dynamic and cluttered displays. Participants searched for a vertical or horizontal line segment among up to 48 other line segments of various orientations, all continuously changing color (from green to red or vice versa). Van der Burg et al. found that search times as well as search slopes were dramatically reduced when the target color change was accompanied by an auditory signal compared to a condition in which the auditory signal was absent (but see Fujisaki, Koene, Arnold, Johnston, & Nishida, 2006). Van der Burg et al. called this benefit the *pip and pop* phenomenon.

Van der Burg et al. (in press) found initial evidence that the pip and pop effect was caused by audiovisual synchrony guiding attention in an automatic, stimulus-driven manner. For instance, search was improved when a target event was accompanied by a tone, even when the

tone was synchronized with a distractor event on 80% of the trials, and thus the tone was not a particularly useful cue. Moreover, in another experiment, search costs were found when the sound occurred simultaneously with distractor events on 100% of the trials and thus never co-occurred with the visual target event.

In this study, we provide converging evidence that audiovisual synchrony in a multiple object environment guides attention in an exogenous fashion by using an experimental setup in which both the auditory and the visual events are completely irrelevant to the task. In Experiments 1 and 2, attentional effects were measured using a temporal order judgment (TOJ) task in which participants were asked to report which of two dots occurred first. In Experiment 3, attentional effects were measured using a simultaneity judgment task (SJ) in which participants were asked to report whether the two dots were presented simultaneously or not. Several studies (see, e.g., Shore, Spence, & Klein, 2001; Stelmach & Herdman, 1991) have provided evidence that the perception of temporal order is influenced by attentional allocation. For instance, when presented simultaneously, attended stimuli were perceived to occur before unattended stimuli. Here we use this phenomenon to assess whether audiovisual synchrony automatically attracts attention. We show that an irrelevant distractor neighboring one of the dots captures attention when an irrelevant tone is synchronized with the color change of that distractor. As a result, the perceived temporal order of the two dots is affected in favor of the synchronized distractor location. Since neither the distractors nor the tones were relevant to the TOJ and SJ tasks, we conclude that synchronized auditory–visual events capture attention.

Experiment 1

Figure 1 provides an example of the displays used in this study (see also the demo for an example trial).

Participants were asked to make a TOJ on which of two dots appeared first, with varying stimulus onset asynchronies (SOAs). Completely irrelevant to the task, at randomly chosen intervals on each side of the display, nine distractor disks continuously changed color (between red and green), one at a time. Importantly, the distractor color change prior to the presentation of the first dot could be accompanied by an irrelevant non-spatial tone. If this auditorily–visually synchronized distractor captures attention, then observers should perceive the dot closest to the synchronized distractor as appearing first, as indicated by measure the point of subjective simultaneity (PSS). To control for potential visual effects of the distractor event prior to the first dot, we also included a condition in which the crucial distractor color change was present, but the tone was absent.
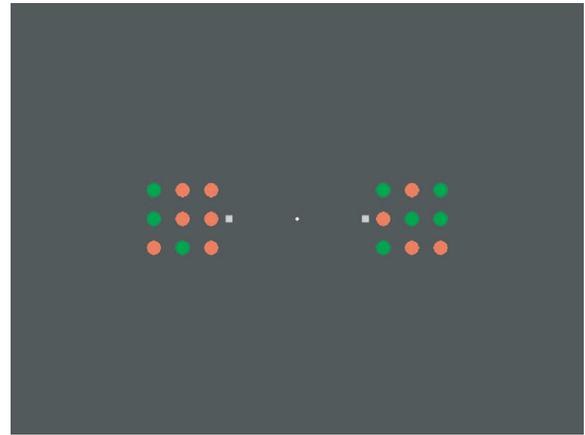

Figure 1. Illustration of the displays used in this study. Two small dots on each side of fixation were used for the temporal order judgment. Participants were asked to report which dot appeared first. Furthermore, 18 irrelevant distractor disks continuously changed color during each trial from red to green or vice versa.

Method

Participants

Twelve students (4 female; mean age 21.5 years; ranging from 18 to 34 years) participated in Experiment 1 as paid volunteers (€7 an hour). All participants were naive as to the purpose of the experiment. Data from one participant were excluded from further analysis because performance was at chance level.

Apparatus and stimuli

Experiments were run in a dimly lit, air-conditioned cabin. Participants were seated at approximately 80 cm from the 19-in. monitor (refresh rate: 120 Hz). The auditory stimulus was a 500-Hz tone (44.1 kHz sample rate; 16 bit; mono) with a duration of 60 ms (including a 5-ms fade-in and fade-out to avoid clicks) presented through Sennheiser HD 202 headphones. The visual stimuli consisted of 18 red (13.9 cd m^{-2}) or green (46.4 cd m^{-2}) distractor disks (radius 0.6°) on a dark gray (4.6 cd m^{-2}) background. Color was randomly determined for each distractor, and half of the disks were placed in an invisible 3 * 3 grid (2.5° * 2.5°) 5.0° to the left of the white (76.7 cd m^{-2}) fixation dot, and the other half of the disks were placed in an identical grid 5.0° to the right of fixation. The two dots subserving the TOJ task were gray (luminance 45.4 cd m^{-2}; 0.3° width; 0.3° height) and positioned 2.9° to the left and to the right of the fixation dot. The display changed 21 times in color during each trial. Each display change was a color change of one randomly selected distractor from green to red or vice versa. One distractor color change was always synchronized with the presentation of a tone (when present), with the constraints that this synchronization

took place at display change 10–15, and that the synchronized distractor was never the one nearest to the gray dot. The duration of the intervals between color changes varied randomly between 50, 100, and 150 ms with the constraint that each interval occurred equally often within each trial, and that the intervals prior to, and after, the presentation of the synchronized distractor were always 150 ms.

Design and procedure

Each trial began with the presentation of a fixation dot for 1,000 ms at the center of the screen, immediately followed by the presentation of the 18 distractor disks. A tone was present on 50% of the trials and synchronized with the color change of the distractor prior to the appearance of the first dot. On the other half of the trials, the tone was absent. On half of the tone-present trials, the tone was synchronized with a distractor on the right side and with a distractor on the left side on the remaining trials. The presentation of the synchronized distractor was always followed by the presentation of the first dot after a fixed interval of 125 ms and subsequently followed by the presentation of the second dot after a randomly determined SOA (−108, −50, −25, −17, −8, 0, 8, 17, 25, 50, and 108 ms). Positive SOAs indicate that the right dot was presented first, and negative SOAs indicate that the left dot was presented first. Tone presence, distractor location, and SOA were mixed within blocks. Participants as well as the distractors were instructed to remain fixated on the fixation dot and to ignore the tone. Participants made an unspeeded response by pressing the z-key when the first dot was presented on the left side or the m-key when the first dot was presented on the right side. There was one practice block of 44 random trials. After the practice block, participants performed 15 experimental blocks of 44 random trials each. Participants received feedback about their overall mean error rate after each block.

Psychometric function fitting

In order to compute the Slope, the PSS, and the JND, the data from each participant were estimated by fitting the following logistic sigmoid function to each individual's data (see also Harrar & Harris, 2005; Spence, Baddeley, Zampini, James, & Shore, 2003) through minimizing the root-mean-square-error (RMSE) in Microsoft Excel Solver:

$$P(\text{response}|\text{SOA}) = \frac{1 - \text{blink rate}}{1 + e^{-\text{slope}(\text{SOA} - \text{PSS})}} \quad (1)$$
$$+ \text{blink rate} \cdot .5,$$

The SOA parameter was fixed and followed the interval between the two dots (−108 to 108 ms). Slope, PSS, and blink rate required estimation. The blink rate parameter

was included to account for a small proportion of trials on which participants were not attending to the actual stimuli (e.g., due to eye blinks or other artifacts), which would otherwise lead to an overestimation of the other parameters (Swanson & Birch, 1992). The blink rate was restricted to a minimum of 0% and a maximum of 2.5% and was estimated to be 0.8% in Experiment 1 and 0.1% in Experiment 2. The slope and PSS were allowed to vary freely. The JND was then estimated from the fitted curve by subtracting the SOA at which the fitted curve crossed the 75% point from the SOA at which the same curve crossed the 25% point and dividing by two.

Results

Percentage "right first" responses

Data from the practice block were excluded from analysis. Figure 2 presents the mean percentage "right first" responses, as a function of SOA (−108 to 108 ms), distractor location (left vs. right), and tone presence (present vs. absent), together with fitted psychometric curves (RMSE = 0.106). These means were subjected to a repeated measures within-subjects univariate ANOVA with the same factors and alpha set at .05. The reported values for p are those after a Huynh–Feldt correction for sphericity violations.

Overall mean percentage "right first" responses was 52.0%. The main effect of SOA was significant, $F(10, 100) = 129.5$, $p < .001$, as participants were complying
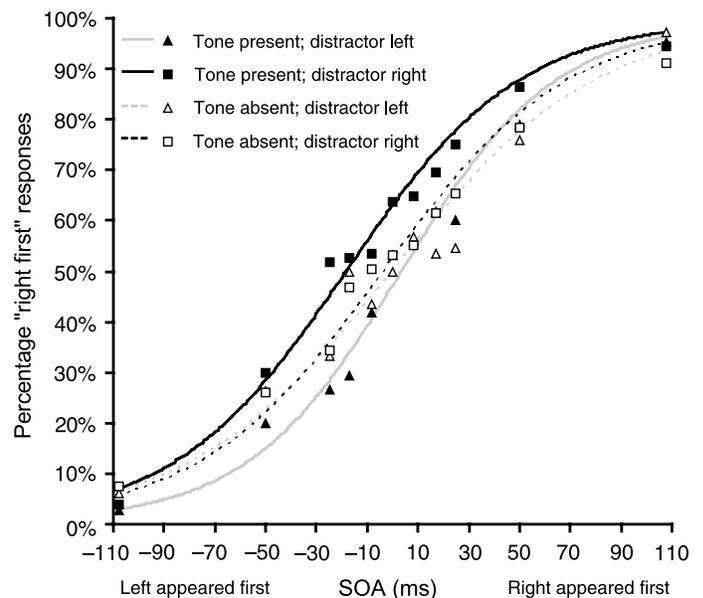

Figure 2. Results of Experiment 1, indicating mean percentage "right first" responses, as a function of SOA, tone presence, and location of the distractor onset, together with fitted psychometric curves. Negative SOAs indicate that the left dot was presented first, positive SOAs indicate that the right dot was presented first.

with the TOJ task. The main effect of tone presence was not significant ($F < 1$). Importantly, the ANOVA revealed a significant two-way interaction between tone presence and distractor location, $F(1, 10) = 12.1$, $p < .01$. This interaction was further examined in detail by two-tailed t tests for each tone-present condition. The t test revealed a significant difference of distractor location when the sound was present, $t(10) = 6.3$, $p < .001$, indicating that overall percentage "right first" responses was greater when the synchronized distractor location was on the right side (58.7%) than when the synchronized distractor location was on the left side (47.8%). There was no significant difference in the sound absent condition, $t(10) = 1.1$, $p = .281$. The interaction between distractor location and SOA approached significance $F(10, 100) = 1.9$, $p = .08$, mostly reflecting the fact that distractor effects were reduced for the longest SOAs (as performance reached ceiling). No other effects were reliable (all p values >.2).

Slope, point of subjective simultaneity (PSS), and just noticeable difference (JND)

Figure 3 represents the extracted parameters from the fitted curves in Figure 2. Slope, PSS, as well as JND were subjected to ANOVAs with tone presence (present and absent) and distractor location (left and right) as within-subject variables and alpha set to .05. Overall mean slope was .029. The ANOVA revealed no significant effects of tone presence, distractor location, and their interaction, $F(1, 10) = 2.4$, $p = .152$, $F < 1$, and $F(1, 10) = 1.2$, $p = .292$, respectively. Overall mean JND was 43.3 ms. The ANOVA revealed no significant effect of distractor location, $F < 1$. There was a trend toward an overall decrease in JNDs when the tone was present, $F(1, 10) = 3.4$, $p = .096$. There was no interaction, $F < 1$.

Overall mean PSS was −4.6 ms. The ANOVA revealed no significant main effect of tone presence on PSS, $F < 1$. Importantly, the main effect of distractor location and the

two-way interaction between tone presence and distractor location were both significant, $F(1, 10) = 13.9$, $p < .005$, and $F(1, 10) = 7.5$, $p < .05$, respectively. The effect of distractor location was further examined by using two-tailed t tests for each tone-present condition. There was a significant effect of distractor location when the tone was present, $t(10) = 6.8$, $p < .001$, but not when the tone was absent, $t < 1$. When the auditory signal was synchronized with a distractor on the left side, the right dot had to lead the left dot by 3.4 ms for simultaneity to be reached, and when the auditory signal was synchronized with a distractor on the right side, the left dot had to lead by 18.3 ms for simultaneity to be reached.

Discussion

The present experiment showed that TOJs were affected when a preceding distractor color change was accompanied by a tone. These effects on TOJs cannot be assigned to the crucial distractor color change prior the TOJ dots alone because TOJs were unaffected when the crucial distractor color change was present, but the tone was absent. Overall, there was a trend toward better performance when a sound was present, suggesting that general alerting may have had a beneficial effect. However, note that alerting cannot explain the shifts in temporal order judgment, as these were dependent on the specific side of the integrated signal.

We see that the overall percentage "right first" responses and PSS measures provides good evidence for the synchronized auditory–visual event biasing attention in the TOJ task. However, we expected performance in the presumably neutral tone-absent conditions to be between the two tone-present conditions. Especially at the longer SOAs, this was not the case. A possible hypothesis is that in the tone-present condition, the tone was partially integrated with the first dot, leading to improved TOJ performance in this condition relative to the tone-absent

Figure 3. Results of Experiment 1: Extracted parameters from the fitted curves in Figure 2. From left to right, the panels show the slope, point of subjective simultaneity (PSS), and just noticeable difference (JND), as a function of tone presence, and distractor location. The error bars represent the .95 confidence intervals for within-subject designs (following Loftus & Masson, 1994) for each specific distractor location/tone presence condition.

condition (which, in addition to general alerting might also explain the trend toward decreased JNDs). To test for this alternative hypothesis, we conducted Experiment 2, in which the tone was always present.

Experiment 2

The present experiment was identical to the previous experiment, except that the two dots to perform the TOJ task were equiluminant with the background in an attempt to make the task more sensitive to attentional manipulations. Furthermore, the tone was always present. On 50% of the trials, a single distractor on the right or left side was accompanied by a tone (as in Experiment 1). On the remaining trials (the control condition), the tone accompanied two simultaneous distractor changes (one on the left side and one on the right side). If auditory–visual synchrony captures attention in an exogenous manner, then we expect to find a bias in TOJs toward the irrelevant distractor in the unilateral distractor condition, relative to the bilateral distractor condition.

Method

Participants

Twelve new students (6 female; mean age 20.0 years; ranging from 18 to 28 years) participated in Experiment 2. All participants were naive as to the purpose of the experiment. Data from two participants were excluded from further analysis because of exceptionally bad performance at the longest SOA (30% incorrect).

The experiment was identical to Experiment 1, except that the two dots subserving the TOJ task were now blue and equiluminant with the background (4.6 cd m^{-2}). Furthermore, the tone was present on all trials. On 50% of the trials, the tone was synchronized with a single distractor color change on either the left or right side. On the remaining 50% (the control condition), the auditory signal was synchronized with two distractor color changes, one on the left side and one on the right side, in mirrored positions. Distractor location (left, right, and both sides) was randomly determined and mixed within blocks. There was one practice block and 15 experimental blocks of 44 randomly selected trials each.

Results

Percentage "right first" responses

Figure 4 presents the mean percentage "right first" responses, as a function of SOA (−108 to 108 ms), and distractor location (left, right, or bilateral), together with fitted psychometric curves (RMSE = 0.077). These data

Figure 4. Results of Experiment 2, indicating mean percentage "right first" responses, as a function of SOA, and location of the synchronized distractor change, together with fitted psychometric curves. Negative SOAs indicate that the left dot was presented first, positive SOAs indicate that the right dot was presented first.

were subjected to a repeated measures Univariate ANOVA with the same factors.

Overall mean percentage "right first" responses was 52.7%. The main effect of SOA was significant, $F(10, 90) = 133.0$, $p < .001$, as participants were complying with the TOJ task. Importantly, the ANOVA revealed a significant effect of distractor location, $F(2, 18) = 23.3$, $p < .001$. As confirmed by separate t tests, mean percentage "right first" responses was lower when the auditory signal was synchronized with a distractor located on the left side (45.4%) than when the auditory signal was synchronized with a distractor located on the right side (59.5%), $t(9) = 5.5$, $p < .001$. Furthermore, percentage "right first" responses was significantly lower in the bilateral control condition (53.2%) than when the auditory signal was synchronized with a distractor located on the right side and significantly higher than when the auditory signal was synchronized with a distractor located on the left side, $t(9) = 3.2$, $p = .01$, and $t(9) = 4.8$, $p = .001$, respectively. The interaction between SOA and distractor location was also significant, $F(20, 180) = 1.9$, $p < .05$. At shorter SOAs, percentage "right first" responses differed considerably for each distractor location, whereas for longer SOAs, they converged to 0% or 100%.

Slope, point of subjective simultaneity (PSS) and just noticeable difference (JND)

Figure 5 presents the extracted parameters from the fitted curves in Figure 4. Slope, PSS, as well as JND were
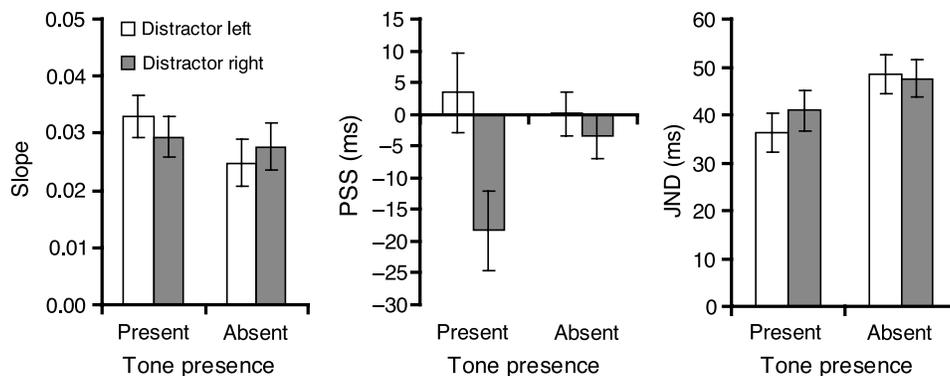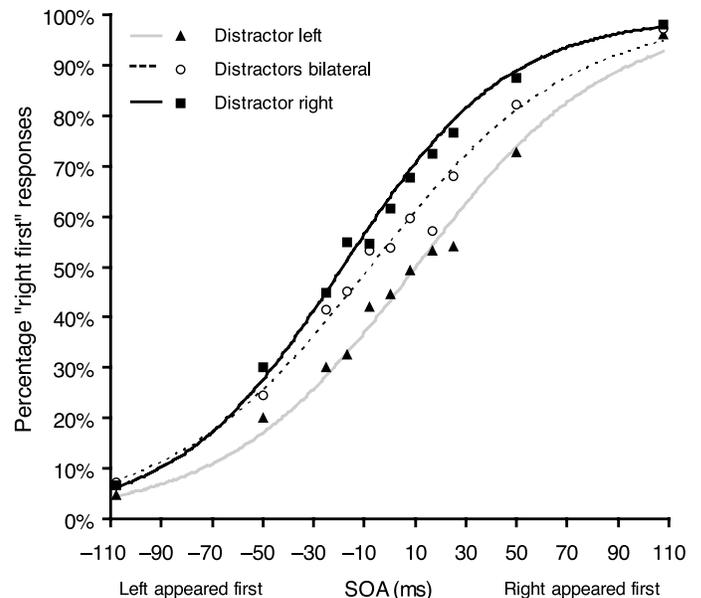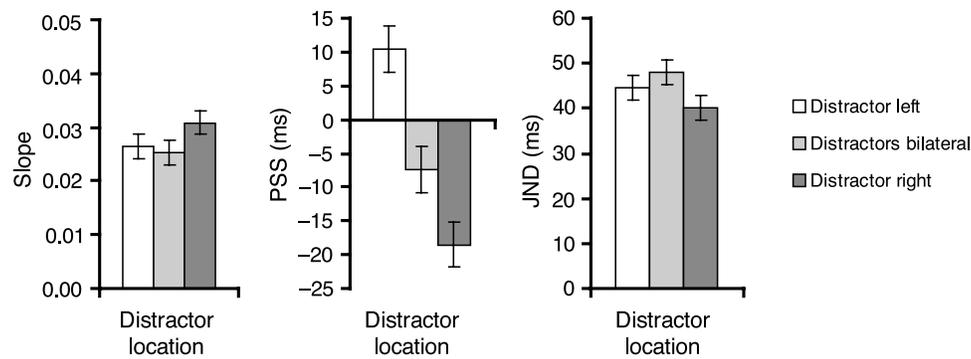
Figure 5. Results of Experiment 2: Extracted parameters from the fitted curves in Figure 4. From left to right, the panels show the slope, point of subjective simultaneity (PSS), and just noticeable difference (JND), as a function of distractor location. The error bars represent the .95 confidence intervals for within-subject designs (following Loftus & Masson, 1994). Here, the confidence intervals reflect those for the main effect of distractor location.

subjected to ANOVAs with distractor location (left, right, or bilateral) as within-subject variables. Overall mean slope was .028. The ANOVA revealed no significant effect of distractor location, $F(2, 18) = 2.9$, $p = .111$. Overall mean JND was 44.3 ms. The ANOVA revealed no significant effect of distractor location on JND, $F(2, 18) = 2.9$, $p = .105$.

Overall mean PSS was −5.1 ms. The ANOVA revealed a significant effect of distractor location on PSS, $F(2, 18) = 24.8$, $p < .001$. Pairwise two-tailed t tests revealed reliable effects on PSS between all conditions [left vs. right, $t(9) = 5.8$, $p < .001$; left vs. bilateral, $t(9) = 4.9$, $p = .001$; and right vs. bilateral, $t(9) = 3.1$, $p = .013$]. When the tone was synchronized with a distractor on the right, the left dot had to lead the right dot by 18.5 ms for simultaneity to be reached. When the tone was synchronized with a distractor on the left, the order was reversed, as the right dot had to lead the left dot by 10.4 ms for simultaneity to be reached. Importantly, in the bilateral control condition, the left dot had to lead the right dot by 7.3 ms for simultaneity to be reached, which was well in between the two unilateral conditions.

Discussion

Consistent with Experiment 1, the present experiment showed that TOJs were affected when a preceding distractor color change was accompanied by a tone. The present experiment provides more evidence against any role of an alerting effect because performance was completely dependent on the location of the synchronized distractor while the tone was always present. Furthermore, results could not be differentially affected by the tone accidentally binding to one of the dots rather than to the distractors since the tone was present in all conditions. We propose that the effects are due to a shift of attention toward the irrelevant distractor, as caused by the integration of the auditory and visual events.

However, it is possible that participants responded directly to the synchronized distractor disk rather than to the first of the two dots when making the TOJ. This would mean that the results are due to a response bias rather than due altered temporal order perception. Note that this account would still imply that observers *perceive* the synchronized distractor (since they now respond to it) despite it being irrelevant to the task, and thus this account would still allow us to conclude that synchronous audiovisual events demand our attention. Nevertheless, our conclusion that the TOJs are affected by the audio–visual event is only justified when we find a case in which such biases can be assumed to be absent, so that attention affects perception rather than response selection. Experiment 3 was designed exactly for that purpose.

Experiment 3

This experiment was identical to Experiment 2, except that participants were asked to judge whether the two dots were presented simultaneously or not (simultaneity judgment, SJ). If auditory–visual synchrony captures attention in an exogenous manner and attention affects SJs, then we expect to find a bias in SJs toward the irrelevant distractor. In contrast, if the results of Experiments 1 and 2 are due to response biases, then we do not expect to find a bias in SJs toward the irrelevant distractor because the synchronized distractor event bears no systematic relationship to the task (Santangelo & Spence, 2008; Schneider & Bavelier, 2003; Zampini, Guest, Shore, & Spence, 2005).

Method
Participants

Eight new students (4 female; mean age 22.8 years; ranging from 21 to 30 years) participated in Experiment 3.

All participants were naive as to the purpose of the experiment. The experiment was identical to Experiment 2, except that the tone was synchronized with a single distractor color change on either the left or right side. Furthermore, participants were asked whether the two dots were presented simultaneously or not by pressing the j- or n-key, respectively. Distractor location (left and right) was randomly determined and mixed within blocks. SOA was also randomly mixed within blocks, with the constraint that there were more simultaneous trials ($N = 300$) than asynchronous trials ($N = 30$ for each SOA) so that a priori the number of synchronous and asynchronous trials was equal (for a similar methodology, see Santangelo & Spence, 2008; Zampini et al., 2005). There was 1 practice block and 15 experimental blocks of 40 randomly selected trials each.

Psychometric function fitting

In order to compute the PSS, the data from each participant were estimated by fitting the following four parameter Gaussian function to each individual's data (see also Santangelo & Spence, 2008) through minimizing the RMSE in Microsoft Excel Solver:

$$P(\text{response}|\text{SOA}) = \text{blink rate} + a \cdot e^{\left[-.5\left(\frac{\text{SOA}-\text{PSS}}{b}\right)^2\right]}, \quad (2)$$

The SOA parameter was equal to the interval between the two dots (-108 to 108 ms). Parameters a and b and blink rate required estimation. The blink rate was restricted to a minimum of 0% and a maximum of 2.5% and was estimated to be 1.3% in Experiment 3. The other parameters were restricted to a minimum of 0.

Results
Percentage "simultaneous" responses

Figure 6 presents the mean percentage "simultaneous" responses, as a function of SOA (-108 to 108 ms), and distractor location (left or right), together with fitted psychometric functions (RMSE = 0.068). These data were subjected to an ANOVA with the same factors.

Overall, the mean percentage "simultaneous" responses was 51.2%. There was a main effect of SOA $F(10, 70) = 55.1$, $p < .001$, as participants were complying with the SJ task. The main effect of distractor location was not reliable ($F < 1$). Importantly, the ANOVA revealed a significant two-way interaction between distractor location and SOA, $F(10, 70) = 3.7$, $p = .001$. This interaction was further examined in detail by two-tailed t tests for negative and positive SOAs. The t test revealed a significant difference of distractor location for negative SOAs (left first; averaged over SOA < 0 ms), $t(7) = 3.0$,
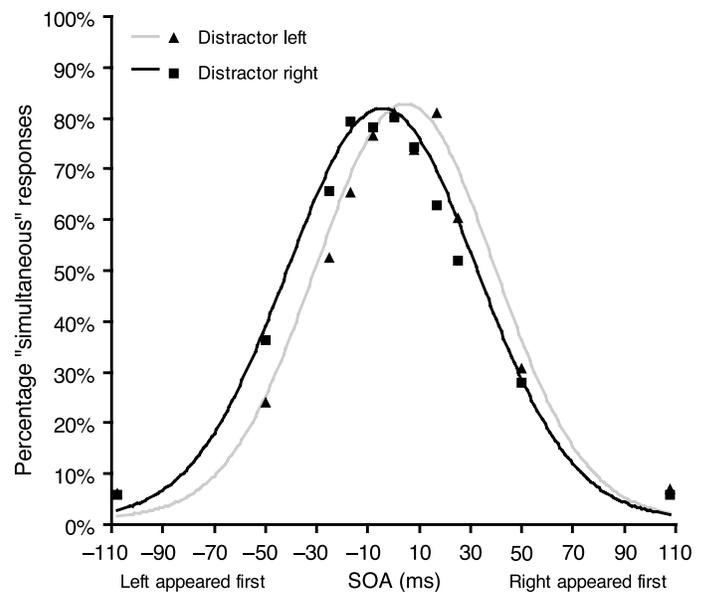

Figure 6. Results of Experiment 3, indicating mean percentage "simultaneous" responses, as a function of SOA, and location of the synchronized distractor change, together with fitted psychometric curves. Negative SOAs indicate that the left dot was presented first, positive SOAs indicate that the right dot was presented first.

$p < .05$, indicating that overall percentage "simultaneous" responses was greater when the synchronized distractor was on the right side (53.1%) than when the synchronized distractor was located on the left side (44.8%). In contrast, for positive SOAs (right first; averaged over SOA > 0 ma), overall percentage "simultaneous" responses was greater when the synchronized distractor was on the left side (50.5%) than when the synchronized distractor was located on the right side (44.5%), $t(7) = 3.7$, $p < .01$.

Point of subjective simultaneity (PSS)

The extracted parameters from the fitted curves in Figure 6 were subjected to separated t tests. Overall mean PSS was 0 ms. The t test revealed a significant effect of PSS on distractor location, $t(7) = 3.1$, $p = .017$, as the PSS was greater when the tone was synchronized with a distractor on the left (4.6 ms) than when the tone was synchronized with a distractor on the right side (-4.4 ms). All other parameters (a, b, and *blink rate*) failed to yield significant differences, all ts < 1.2.

Discussion

The present pattern of results is consistent with that found in Experiments 1 and 2: A synchronized distractor affected judgments on whether one of two dots appeared

first or whether they appeared simultaneously. Using SJs rather than TOJS, the experiment rules out a response bias as the sole explanation of the results obtained in Experiments 1 and 2 because SJs are assumed to be insensitive to responses biases (Santangelo & Spence, 2008; Schneider & Bavelier, 2003; Zampini et al., 2005). Instead, we propose that the effects are due to a shift of attention toward the irrelevant distractor, as caused by the integration of the auditory and visual events. It deserves mentioning, however, that the present experiment revealed a smaller PSS than found in Experiments 1 and 2 (9.0 vs. 21.7 and 28.9 ms, respectively), and one might argue that the shifts in PSS in the latter experiments were at least partly due to response biases. As we argued earlier, even such a response bias would be a sign of participants failing to ignore the synchronized event.

General discussion

The present study demonstrates that visual TOJs and SJs were affected when a preceding distractor color change was accompanied by a spatially non-specific auditory signal (Experiments 1, 2, and 3). In other words, an irrelevant tone synchronously presented with an irrelevant visual event affected spatial processing in a multiple object environment. Furthermore, the present findings are neither due to a visual effect (Experiment 1) nor due to the auditory signal acting as a temporal cue, or alerting signal (Experiments 1 and 2). Moreover, Experiment 3 ruled out an explanation in terms of response bias. The present findings are consistent with the idea that audiovisual synchrony guides attention in an exogenous manner—a phenomenon that we have dubbed "pip and pop" in previous work (Van der Burg et al., in press).

The present study is not the first to show effects of non-spatial auditory signals on visual temporal perception. Other studies have shown TOJs to benefit when the first dot was preceded by a tone while the second dot was followed by a tone. This auditory stretching of the interval between visual events is known as *temporal ventriloquism* (see, e.g., Morein-Zamir, Soto-Faraco, & Kingstone, 2003; Vroomen & Keetels, 2006). In these studies, the perceived temporal properties of the visual event, such as its onset time and duration, were attracted toward those of a sound with which the visual stimulus was associated. Whereas these studies have shown TOJs to benefit from presenting tones, this study revealed impaired TOJs and SJs as well (in the sense that the point of subjective simultaneity does not correspond to the point of physical simultaneity). Thus, even though we presented a tone prior to the two dots, this tone did not lead to an improved TOJ or SJ per se. Moreover, Morein-Zamir et al. (2003) have shown that temporal ventriloquism is due to the second tone trailing the second dot and not due to the first

tone acting as a warning or alerting signal. Therefore, with regard to this study, the effect on TOJs and SJs cannot be attributed to temporal ventriloquism.

By showing automatic integration in multiple object environments, our findings extend earlier work on temporal integration of single auditory and visual events at a time (Olivers & Van der Burg, 2008; Vroomen & De Gelder, 2000). The findings are also consistent with Santagelo and Spence (2007) who have shown that only bimodal spatial cues are able to capture spatial attention when observers are performing a demanding task. Santangelo and Spence argued for increased perceptual saliency of bimodal cues relative to unimodal cues. Whereas in Santangelo and Spence's study the auditory cue was spatially localized, this study shows that the auditory event need not be localizable for it boost the saliency of a visual event (though see our next point). Moreover, several neurological studies have shown that auditory signals can activate early "unisensory" visual areas (see, e.g., Giard & Peronnet, 1999; Molholm et al., 2002; Talsma, Doty, & Woldorff, 2007). This early (~40–50 ms post-stimulus onset) activation supports the notion that visual processing is modified by auditory inputs well before it is completed. With regard to this study, we propose that the auditory signal modulates processing within the visual cortex, allowing the auditory signal to interact with the irrelevant distractor color change. As a result, this auditory modulation increases the visual saliency of the distractor which then attracts attention. Future research will have to establish what the exact nature of this auditory–visual modulation is. For example, the auditory signal may boost the gain of visual neurons, or it may increase baseline firing rates.

Although we wish to frame our results in terms of the auditory signal directly boosting the visual signal, an alternative, more indirect explanation cannot be excluded. Such an explanation would involve *spatial ventriloquism*—the phenomenon that the perceived location of sound is shifted toward that of a synchronous visual event (Slutsky & Recanzone, 2001; Spence & Driver, 2000; Thomas, 1941; Vroomen, Bertelson, & De Gelder, 2001). According to this account, participants may have perceived the spatially non-specific auditory signal as emanating from the synchronized distractor's location. The now localizable sound attracts attention toward the visual locations in an automatic, exogenous fashion (Spence & Driver, 2000; Vroomen et al., 2001), which then results in shifted TOJs and SJs. Thus, the effects would be due to auditory attentional capture rather than visual attentional capture. This remains an important question for the future. So far, unlike this study, studies investigating the ventriloquism illusion used a single visual event at a time, which was highly salient in its environment (Slutsky & Recanzone, 2001; Spence & Driver, 2000; Thomas, 1941; Vroomen et al., 2001). Moreover, studies have shown that ventriloquism is only present when the distance between the auditory and the

visual stimuli is relatively small (Slutsky & Recanzone, 2001). New studies would therefore probably benefit from investigating the role of spatial separation between the auditory and the visual sources on the pip and pop effect. Nevertheless, whether the underlying mechanism consists of audition modulating vision or of vision modulating audition, the present work shows that audiovisual integration automatically guides attention, biasing the competition between objects in cluttered, dynamic displays.

Acknowledgments

Corresponding author: Erik van der Burg.
Email: e.van.der.burg@psy.vu.nl.
Address: Van der Boechorststraat 1, 1081 BT, Amsterdam, The Netherlands.

References

Bertelson, P., & Radeau, M. (1981). Cross-modal bias and perceptual fusion with auditory-visual spatial discordance. *Perception & Psychophysics, 29,* 578–584. [PubMed]

Doyle, M. C., & Snowden, R. J. (2001). Identification of visual stimuli is improved by accompanying auditory stimuli: The role of eye movements and sound location. *Perception, 30,* 795–810. [PubMed]

Fujisaki, W., Koene, A., Arnold, D., Johnston, A., & Nishida, S. (2006). Visual search for a target changing in synchrony with an auditory signal. *Proceedings of the Royal Society B: Biological Sciences, 273,* 865–874. [PubMed] [Article]

Giard, M. H., & Peronnet, F. (1999). Auditory–visual integration during multimodal object recognition in humans: A behavioral and electrophysiological study. *Journal of Cognitive Neuroscience, 11,* 473–490. [PubMed]

Harrar, V., & Harris, L. R. (2005). Simultaneity constancy: Detecting events with touch and vision. *Experimental Brain Research, 166,* 465–473. [PubMed]

Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review, 1,* 476–490.

Molholm, S., Ritter, W., Murray, M. M., Javitt, D. C., Schroeder, C. E., & Foxe, J. J. (2002). Multisensory auditory-visual interactions during early sensory processing in humans: A high-density electrical mapping study. *Cognitive Brain Research, 14,* 115–128. [PubMed]

Morein-Zamir, S., Soto-Faraco, S., & Kingstone, A. (2003). Auditory capture of vision: Examining temporal ventriloquism. *Cognitive Brain Research, 17,* 154–163. [PubMed]

Olivers, C. N., & Van der Burg, E. (2008). Bleeping you out of the blink: Sound saves vision from oblivion. *Brain Research.* [PubMed]

Santangelo, V., & Spence, C. (2007). Multisensory cues capture spatial attention regardless of perceptual load. *Journal of Experimental Psychology: Human Perception and Performance, 33,* 1311–1321. [PubMed]

Santangelo, V., & Spence, C. (2008). Crossmodal attentional capture in an unspeeded simultaneity judgment task. *Visual Cognition, 16,* 155–165.

Schneider, K. A., & Bavelier, D. (2003). Components of visual prior entry. *Cognitive Psychology, 47,* 333–366. [PubMed]

Shore, D. I., Spence, C., & Klein, R. M. (2001). Visual prior entry. *Psychological Science, 12,* 205–212. [PubMed]

Slutsky, D. A., & Recanzone, G. H. (2001). Temporal and spatial dependency of the ventriloquism effect. *Neuroreport, 12,* 7–10. [PubMed]

Spence, C. (2007). Audiovisual multisensory integration. *Acoustical Science and Technology, 28,* 61–70.

Spence, C., Baddeley, R., Zampini, M., James, R., & Shore, D. I. (2003). Multisensory temporal order judgments: When two locations are better than one. *Perception & Psychophysics, 65,* 318–328. [PubMed]

Spence, C., & Driver, J. (2000). Attracting attention to the illusory location of a sound: Reflexive crossmodal orienting and ventriloquism. *Neuroreport, 11,* 2057–2061. [PubMed]

Stelmach, L. B., & Herdman, C. M. (1991). Directed attention and perception of temporal order. *Journal of Experimental Psychology: Human Perception and Performance, 17,* 539–550. [PubMed]

Swanson, W. H., & Birch, E. E. (1992). Extracting thresholds from noisy psychophysical data. *Perception & Psychophysics, 51,* 409–422. [Pubmed]

Talsma, D., Doty, T. J., & Woldorff, M. G. (2007). Selective attention and audiovisual integration: Is attending to both modalities a prerequisite for early integration? *Cerebral Cortex, 17,* 679–690. [PubMed]

Thomas, G. J. (1941). Experimental study of the influence of vision on sound localization. *Journal of Experimental Psychology, 28,* 167–177.

Van der Burg, E., Olivers, C. N. L., Bronkhorst, A. W., & Theeuwes, J. (in press). Pip and pop: Non-spatial auditory signals improve spatial visual search. *Journal of Experimental Psychology: Human Perception and Performance*.

Vroomen, J., Bertelson, P., & de Gelder, B. (2001). Directing spatial attention towards the illusory location of a ventriloquized sound. *Acta Psychologica, 108,* 21–33. [PubMed]

Vroomen, J., & de Gelder, B. (2000). Sound enhances visual perception: Cross-modal effects of auditory organization on vision. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 1583–1590. [PubMed]

Vroomen, J., & Keetels, M. (2006). The spatial constraint in intersensory pairing: No role in temporal ventriloquism. *Journal of Experimental Psychology: Human Perception and Performance, 32,* 1063–1071. [PubMed]

Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin, 88,* 638–667. [PubMed]

Zampini, M., Guest, S., Shore, D. I., & Spence, C. (2005). Audio-visual simultaneity judgments. *Perception & Psychophysics, 67,* 531–544. [PubMed]