Does Number of Perceptions or Cross-Modal Auditory Cueing Influence Audiovisual Processing Speed?

Author(s): Nicholas Altieri, Michael J. Wenger, Mark T. Wallace and Ryan A. Stevenson

Source: *The American Journal of Psychology*, Vol. 129, No. 1 (Spring 2016), pp. 11-21

Published by: University of Illinois Press

Stable URL: https://www.jstor.org/stable/10.5406/amerjpsyc.129.1.0011

Does Number of Perceptions or Cross-Modal Auditory Cueing Influence Audiovisual Processing Speed?

NICHOLAS ALTIERI
Idaho State University

MICHAEL J. WENGER
The University of Oklahoma

MARK T. WALLACE
Vanderbilt University

RYAN A. STEVENSON
University of Toronto

What factors contribute to redundant target processing speed besides statistical facilitation? One possibility is that multiple percepts may drive these effects. Another, although not mutually exclusive hypothesis, is that cross-channel cueing from one modality to another may influence response times. We implemented an auditory–visual detection task using the sound-induced flash illusion to examine whether one or both of these possibilities contributes to changes in processing speed; we did so by examining the data of individual participants. Our results indicated shorter response times in several participants when multiple flashes were perceived in the standard sound-induced flash illusion, thereby replicating previous work in the literature. Additionally, we found evidence for faster responses in several participants when carrying out the same analysis in trials in which 1 beep was presented with 2 real flashes. Overall, our analysis indicates that some observers benefit from cross-modal facilitation, whereas others may benefit from a combination of cross-modal facilitation and increased perceptual judgments.

KEYWORDS: cross-modal facilitation, cross-channel cueing

Research on audiovisual perception has shown that responses to multiple targets are often faster (e.g., Miller, 1982, 1986) than the fastest of the displays containing single targets (Grice, Canham, & Gwynne, 1984). Consider, for example, a simple auditory–visual detection task requiring participants to make a button-press response as soon as they see or hear a target. In experiments such as these, response times (RTs) in the auditory plus visual signal condition are typically shorter than in either of the single-target conditions. Examples of such facilitation have also occurred in the context of visual detection (Mordkoff

American Journal of Psychology
Spring 2016, Vol. 129, No. 1 pp. 11–21 •

& Yantis, 1991; Schroter, Fiedler, Miller, & Ulrich, 2011; Townsend & Nozawa, 1995), auditory detection (Schroter, Frei, Ulrich, & Miller, 2009), and even for audiovisual tasks involving speech signals (Altieri & Townsend, 2011; Altieri & Wenger, 2013).

The benefit of having multiple signals available has often been explained in terms of statistical facilitation (Raab, 1962). Historically, the model-based framework used to explain such facilitation is the independent race model, or simply called race model (Miller, 1982). In this framework, auditory or visual inputs simultaneously accrue evidence toward a detection threshold. As more items are added to the display, recognition occurs more quickly. This phenomenon is simply a statistical result of having more independent items in the display, which increases the chance that detection occurs earlier because processing may cease as soon as one item surpasses threshold and "wins the race." Under certain circumstances, however, the degree of redundant target influence may be too great to be explained by statistical facilitation alone. Such race model violations have been explained by coactive processes in which the activations in each channel are combined (Giray & Ulrich, 1993; Miller, 1982, 1986; Townsend & Nozawa, 1995) or interact with one another at various stages in sensory processing (Eidels, Houpt, Altieri, Pei, & Townsend, 2011; Otto & Mamassian, 2012; Savazzi & Marzi, 2008; Townsend & Altieri, 2012; Townsend & Wenger, 2004).

Effects of Multiple Percepts

Recent investigations have refined our knowledge of the factors that may lead to the "redundant signals" effects and multimodal processing speed in general: Results have suggested that the number of percepts may actually drive the speed of redundant target processing rather than the number of items in the display per se (Schroter et al., 2009, 2011; Ulrich, Miller, & Schroter, 2007). In support of this hypothesis, Schroter et al. (2011) demonstrated that stereoscopically fused stimuli (two visual flashes perceived as one) failed to elicit redundant target facilitation, whereas two flashes presented at noncorresponding areas of the retina (two flashes perceived as two flashes) yielded the predicted facilitation.

These unisensory visual findings have also been extended to show that visual detection speed can be increased when presented in conjunction with multiple auditory percepts that did not function as targets. This can be demonstrated with the sound-induced flash illusion (Shams, Kamitani, & Shimojo, 2000). In this illusion, when a single flash is accompanied by two auditory tones in close temporal synchrony, the perception of an illusory double flash is often elicited. Perceiving this illusory double flash has been shown to elicit similar increases in visual detection speed, as does presenting a true double flash (Fiedler, O'Sullivan, Schroter, Miller, & Ulrich, 2011). Findings such as these suggest that redundant target facilitation may be driven by the number of perceptions rather than the actual number of targets.

Alternatively, depending on the stimuli used in the task, information from one target or modality may provide facilitatory or inhibitory information to the other. This may lead to a speeding up or slowing down of redundant target processing (Altieri & Wenger, 2013; Eidels et al., 2011).

The question investigated in this report, then, concerns whether changes in processing speed associated with the sound-induced flash illusion are driven by the number of percepts, facilitation or inhibition across modalities, or both. Each framework is depicted in Figure 1. According to the cross-modal facilitation theory, for example, the double beep in the auditory modality increases visual processing speed.

Our study implemented a full-factorial design in an audiovisual detection study to assess whether one or perhaps both of these frameworks hold by examining the data of individual participants. In our design we sought to reproduce the "fission effect," in which an illusory flash was presented with two beeps. We also sought to extend and elaborate on these findings by exploring changes in processing speed under conditions in which a single beep is paired with two flashes. This latter phenomenon has been reported to drive the perception of single flashes (known as the "fusion effect"; see Andersen, Tiippana, & Sams, 2004; Mishra, Martinez, & Hillyard, 2008; Shams, Ma, & Beierhold, 2005).

The following study required participants to make a "yes" response when an auditory or visual target was perceived and subsequently perform a perceptual judgment indicating the number of flashes they perceived. The target could be a single or multiple stimuli in the auditory, visual, or both modalities. The audiovisual stimulus structure contained the following trial types: one beep + one flash (A1V1), two beeps + one flash (A2V1), one beep + two flashes (A1V2), two beeps + two flashes (A2V2).

Frameworks and Hypotheses

NUMBER OF PERCEPTS (THEORY 1).

According to this theoretical framework (Figure 1, top panel), the perception of multiple targets provides a sufficient condition for facilitating multisensory processing speed. Therefore, the main prediction here is that the number of percepts should drive audiovisual processing speed; importantly, this should be true even for participants who do not experience the fusion or fission illusions. Second, we predict faster processing of multiple percepts in the unisensory condition; that is, we predict faster responses in the V2 condition when two flashes were perceived as opposed to only one. Third, because differences are predicted for V2 (flash-only condition) for one versus two perceived flashes, it makes sense that a significant increase in processing speed should emerge between A1V2 (when two flashes are perceived) and A1V1. Fourth, we predict a significant decrease in processing speed for A2V1, when one perceived flash is perceived, versus A2V2, when two flashes are perceived. Overall, this framework predicts a global effect such that faster processing results when a greater number of stimuli are perceived.

CROSS-MODAL FACILITATION (THEORY 2).

An alternative framework (Figure 1, bottom panel) is that processing speed is driven by differential cross-modal activation strengths from the auditory to visual modality (Eidels et al., 2011; Townsend & Wenger, 2004) rather than the number of auditory or visual percepts. The primary prediction here is that

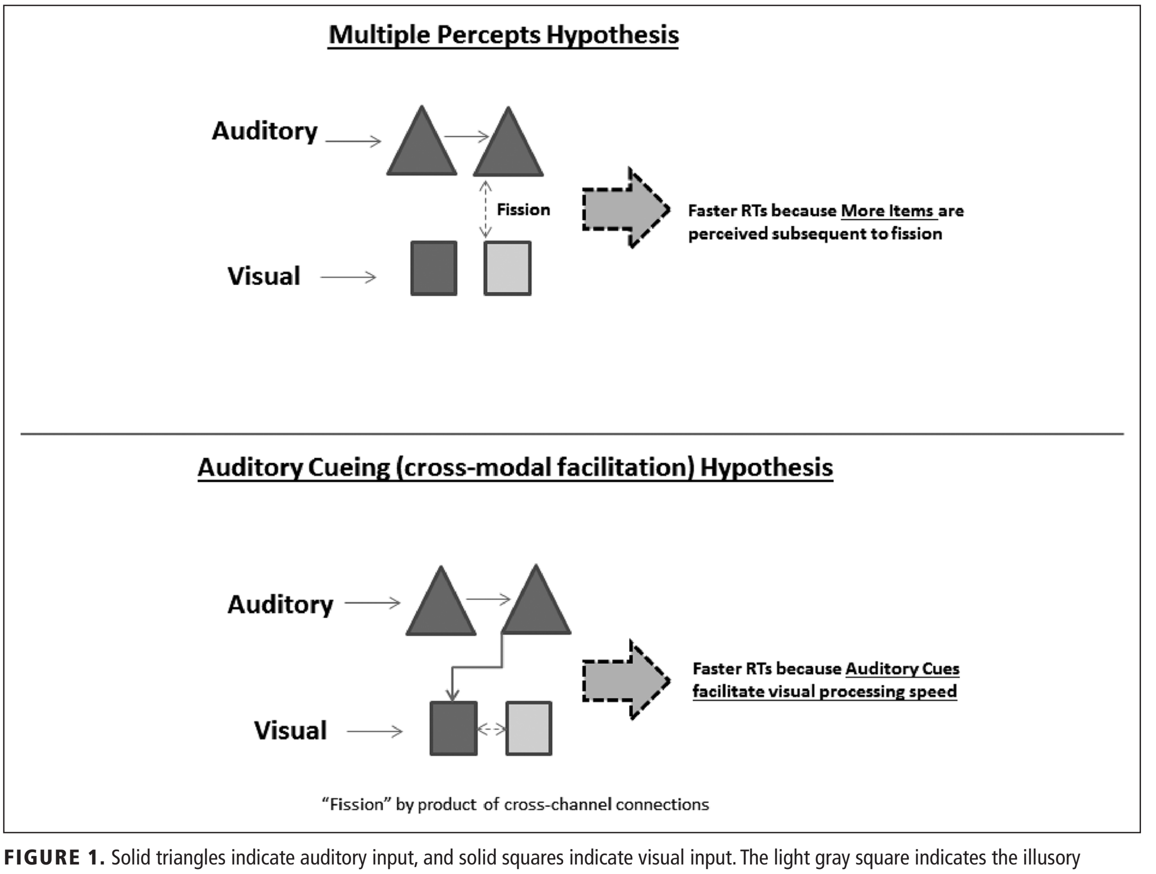

FIGURE 1. Solid triangles indicate auditory input, and solid squares indicate visual input. The light gray square indicates the illusory double percept. The multiple percepts theory is shown in the top panel. This theory posits that RTs are shorter in the presence of the illusory flash simply because more items are perceived. The auditory cueing hypothesis (bottom panel) assumes that auditory cues speed up visual processing via cross-modal interaction; this is exemplified by the arrow

facilitation or inhibition in audiovisual processing speed should be observed only in participants who experience the fission or fusion illusion. That is, we predict that the number of auditory percepts should interact with the number of visual percepts and influence audiovisual processing speed. (It is possible for a participant to experience an illusion but not necessarily benefit or be inhibited in terms of processing speed by the auditory effect.)

There are three possible outcomes for participants who show significant effects for any of the proposed tests:

> Evidence for Theory 1: First, suppose the participant fails to perceive the fission or fusion illusion. However, they show a significant change in audiovisual processing speed for A2V1 or A1V2 (when two vs. one flash is perceived). Observers also show significant results for at least one of the three predictions for Theory 1. Here, we propose that the number of percepts probably drives the effects.
>
> Evidence for Theory 2: We propose that the participant must show evidence for the fission or fusion illusion (interaction between auditory cues and number of visual percepts) and also show the predicted change in audiovisual processing speed for A2V1 or A1V2 when two flashes versus one flash is perceived. Additionally, we predict null results for the last three proposed tests for Theory 1.

Both theories can be supported, and therefore the cause of the effects would be indeterminate if the following occurs: The participant shows evidence for the fission or fusion illusion while also showing the predicted change in audiovisual processing speed for A2V1 or A1V2 (when two vs. one flash is perceived). Additionally, significant results should be observed for at least one of the proposed tests for Theory 1.

METHODS

Participants

Eight right-handed college aged adults (4 women) with normal or corrected-to-normal vision and without history of neurological or hearing impairment (bilateral pure-tone thresholds lower than 25 dB) participated. Participants were recruited from the Idaho State University campus and paid $10 per hour for their participation. The Idaho State University Institutional Review Board approved this study.

Stimuli and Apparatus

Experimental stimuli were presented in E-Prime 2.0 (Psychology Software Tools Inc., Sharpsburg, PA) on a Dell PC computer. Participants were seated in a darkened room approximately 20–24 inches from the center of the computer monitor. The auditory stimuli consisted of 500-Hz pure tones (duration = 20 ms) presented over Beyer-Dynamic 100 Headphones. The tones were played at a comfortable listening volume of approximately 68 dB sound pressure level. The visual flashes (duration ≈ 20 ms) were presented against a black background on a LED Dell flat-screen monitor (Dell P2213 DVI, resolution 1,400 × 1,050, 60 Hz). The flash was an opaque white disk on a black background with a Michelson contrast of $C_M = (21 - 1)/(21 + 1) = .91$. The approximate diameter of the disk was 1° visual angle. On each trial, the visual stimulus was flashed 10° above or below the central fixation point. Each trial began with a centrally positioned white + sign (0.5° × 0.5° visual angle), indicating that the trial was about to begin while also serving as a fixation point. Participants used a mouse to initiate trials (left button) and for collecting response accuracy and RTs (right button).

Procedure

The experiment involved two separate sessions consisting of four blocks each, with an optional rest period between blocks. The sessions were separated by one day at most. The experiment began with 48 practice trials to acquaint the participant with the basic protocol and go–no-go procedure. Each block consisted of 200 randomly ordered trials, with each participant receiving a total of 1,600 trials. Probabilistic stimulus configurations that could bias multisensory target responses were addressed by balancing audiovisual, auditory/visual-only, and target-absent trials, as in Schroter et al. (2011) (cf. Mordkoff & Yantis, 1991, for discussion on contingencies). The probability of either an auditory tone or a visual flash (independently of whether the other was present) was set to $p = .60$. Therefore, "go" responses were required in 84% (1,344) of trials, and the remaining 16% consisted of target-absent (256) catch trials, which required observers to withhold their response. In 576 go-trials, participants were presented with simultaneous beeps and flashes in the following configurations: 144-single beep with one flash (A1V1), 144-double beep with one flash (A2V1), 144-single beep with two flashes (A1V2), and 144-double beep with two flashes (A2V2). In 768 single-modality "go" trials, observers completed 384 trials in which they were presented with visual-only stimuli (192 single-flash [V1] and 192 double-flash [V2]) and 384 trials in which they were presented with auditory-only stimuli (192 single-beep [A1] and 192 double-beep [A2]).

At the beginning of each trial, a white dot appeared in the center of the screen. This was the cue for participants to initiate the trial by using their left index finger to press the left button on the mouse. A white fixation cross (+) appeared in the center of the screen, and participants were instructed to fixate on the cross throughout the trial. On target-present go-trials, the stimulus (A-only, V-only, or AV) appeared after a randomized delay period of 400–700 ms (uniform distribution). Randomized intervals were used between fixation and stimulus onset in an effort to reduce anticipatory responses. Participants were instructed to make a speeded response as soon as they perceived the stimulus. In the audiovisual conditions, the first visual stimulus (flash) was always presented simultaneously with the first auditory stimulus (beep). On trials involving either two beeps or two flashes, the second stimulus followed after an asynchrony of 20 ms (cf. Elze, 2010, for use of the term *stimulus onset asynchrony*). The visual stimulus presentation of 20 ms, followed by a stimulus onset asynchrony of 20 ms and a subsequent stimulus of 20 ms, yields an approximate trial time of 60 ms total when presented via an LED monitor with a refresh rate of 60 Hz (Elze, 2010). In our study, the second flash always occurred in the same spatial location as the first. Presentation times were altered slightly from Shams et al. (2000) to help maintain consistent trial length across all unisensory and audiovisual target conditions. Participants were allotted 2,000 ms to make their responses, and they subsequently indicated the number of perceived flashes by using the keyboard to type "1" or "2" into a dialog box after their initial detection response on each trial (not shown in Figure 1). On target-absent catch trials, the fixation cross remained on the screen for 2,000 ms past the random delay period, and participants were instructed to withhold their response. Participants were encouraged to take a short rest between blocks, and they initiated the next set of trials by pressing the mouse button. Figure 2 shows the basic setup

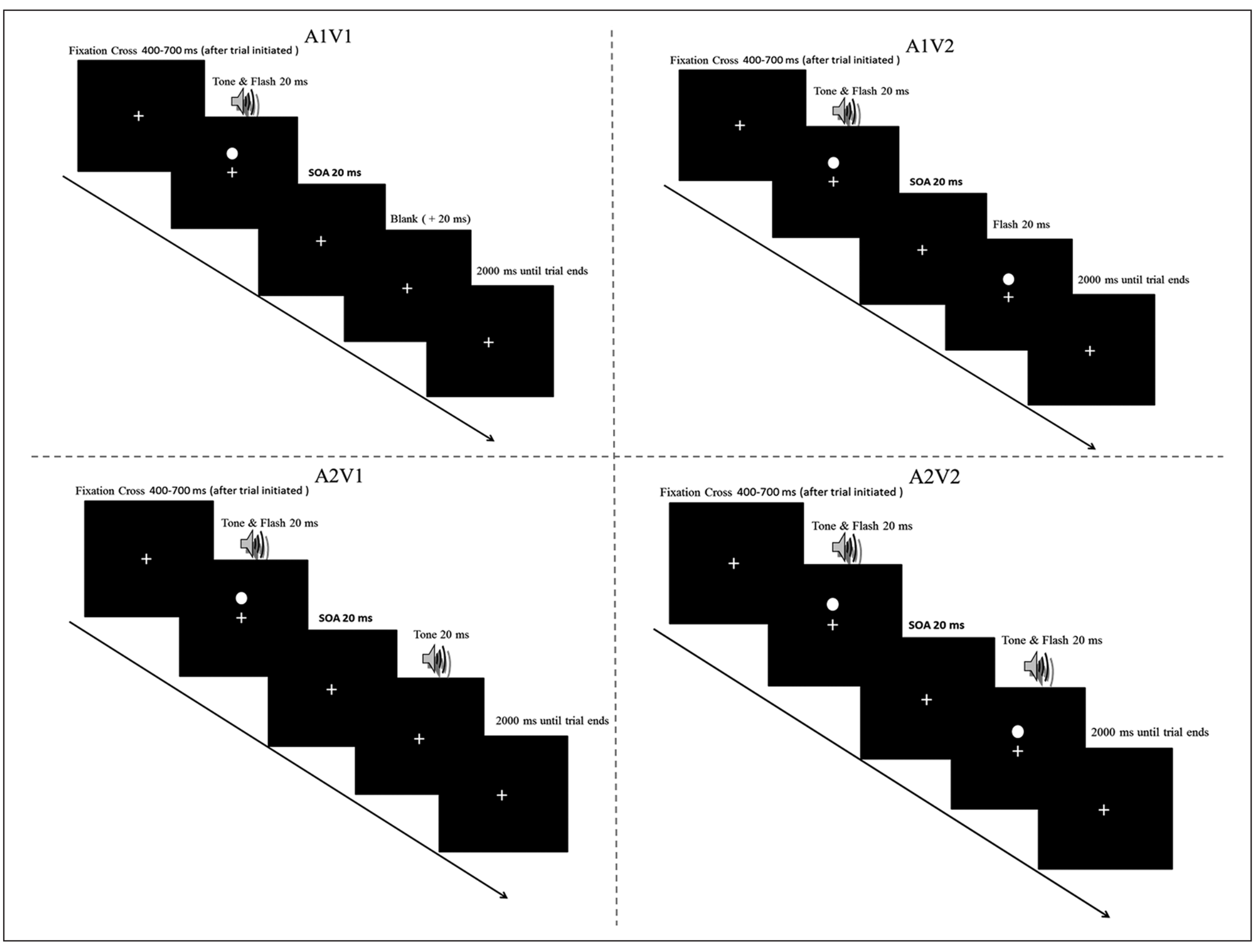

FIGURE 2. Trial time frame for each of the 4 audiovisual conditions (flash above fixation)

of each of the audiovisual experimental conditions at initiation of each trial. The sequence is similar for unisensory conditions.

RESULTS AND DISCUSSION

Group-Level Analyses

MEAN ACCURACY AND RT.
Participants were accurate in detecting the presence or absence of a target. Across all four audiovisual conditions, mean accuracy was 98.9% correct (SE = 0.21%), and for the unisensory conditions the mean accuracy was 97.8% (SE = 0.10%). The rate of false alarms in all conditions was only 1.3% (SE = 0.64%). Table 1 summarizes the mean RTs averaged across participants for each of the experimental conditions. RTs shorter than 100 ms or less than three standard deviations from the mean were removed from the analyses (less than 2.5% of the data). Only correct RTs were included in our analyses. First, to determine whether processing speed differed across experimental conditions, we carried out a repeated-measures ANOVA; results showed a significant difference in mean RTs, as expected, $F(3, 21) = 7.87$, $p < .001$, $\eta_p^2 = .53$.

In addition to reporting the presence or absence of a target via mouse button press, participants were instructed to perform a judgment indicating whether they saw one or two flashes immediately after each trial in which a flash occurred (piloting indicated that participants rarely if ever observed flashes on auditory-only or target-absent trials). We reported the percentage of trials in which a second flash was observed in Table 1 along with the standard error.

We used paired-samples t tests to investigate the ability of the illusory condition to induce the perception of a second flash relative to physically presenting a second flash. First, when we compared the A2V1 with the V1 condition to test for presence of the fission effect, participants perceived two flashes at a higher rate in the A2V1 condition, as predicted, $t(7) = 5.24, p < .001$. Concordant with prior work, this illustrates that the presence of two beeps when paired with a single flash often results in the perception of two flashes, with one of those flashes being illusory. More appropriately, the binomial sign test comparing whether two flashes were reported at a higher rate in the A2V1 compared with the V1 condition confirmed that each observer showed this effect. Binomial sign test results are shown in Table 2.

In addition to examining whether the fission effect occurred, our design also allowed an assessment of the reciprocal fusion effect, in which the combination of a single beep with multiple flashes has been reported to result in the perception of only a single flash. However, on comparison of the number of perceived flashes in the A1V2 and V2 conditions, we failed to observe a significant difference in the percentage of perceived double flashes at the group level, $t(7) = 1.77$, $p = .12$. Crucially, the binomial sign test showed that five of eight observers showed evidence for the fusion effect with a lower probability of two flashes reported in the A1V2 condition compared with the V2 (see Table 2). One possibility is that participants who experience the fission effect may respond faster to perceived double flashes compared with when they perceive only one flash. In other words, processing strategies may differ for individuals depending on their susceptibility to the fission effect.

TABLE 1. Mean Response Time (*SE*), in ms, for Responses Averaged Across Conditions

Audiovisual	Response time (*SE*)	2 Flashes	Unisensory	Response time (*SE*)	2 Flashes
A1V1	403 (91)	1.33 (0.18)	A1	415 (85)	n/a
A1V2	443 (88)	48.25 (4.53)	A2	410 (89)	n/a
A2V1	387 (89)	50.63 (3.18)	V1	466 (84)	3.75 (0.39)
A2V2	396 (89)	89.00 (1.41)	V2	446 (82)	57.40 (4.42)

Note. *The mean percentage of the time 2 flashes were reported (*SE*) is also displayed. A1V1 = 1 beep + 1 flash; A2V1 = 2 beeps + 1 flash; A1V2 = 1 beep + 2 flashes; A2V2 = 2 beeps + 2 flashes; n/a = not applicable.*

TABLE 2. Binomial Statistical Test Results Comparing the Number of Perceived Flashes Across Trial Types

Participant	A2V1 vs. V1	A1V2 vs. V2
1	$p < .0001^*$	$p < .0001^*$
2	$p < .0001^*$	$p < .0001^*$
3	$p < .0001^*$	$p = .99$
4	$p < .0001^*$	$p = .98$
5	$p < .0001^*$	$p < .01^*$
6	$p < .0001^*$	$p < .0001^*$
7	$p < .0001^*$	$p = .22$
8	$p < .0001^*$	$p < .01^*$

Note. *First, A2V1 vs. V1 tested whether flashes occurred with a higher probability in the A2V1 trials, and A1V2 vs. V2 tested whether flashes occurred at a significantly higher rate in V2 trials.* $^*p < .05$.

CHANGES IN RT AS A FUNCTION OF PERCEPTUAL REPORTS.
To investigate changes in RT as a function of perceptual report, we carried out a repeated-measures ANOVAs with condition (A1V2 vs. A2V1) and perceived flashes (1 vs. 2) as factors. (Conditions A1V1 and A2V2 were not included because two flashes were rarely reported for A1V1, and one flash was rarely reported for A2V2 trial types.) Results indicated significantly faster responses when two flashes were reported, $F(1, 14) = 12.26$, $p < .01$, $\eta_p^2 = .47$; however, the main effect for condition was nonsignificant, $F(1, 14) < 1$, $\eta_p^2 = .01$, as was the interaction, $F(1, 14) < 1$, $\eta_p^2 = .02$.

To further investigate the results, we carried out a paired-sample *t* test comparing mean RTs using trials from the A2V1 condition when two flashes were reported with cases in which only one flash was reported. Results showed faster responses when two flashes were reported, $M = 382$ ms, $SE = 20$ ms; one flash, $M = 409$ ms, $SE = 23$ ms, $t(7) = 4.21$, $p = .006$. Next, we found evidence for marginally longer RTs when one flash was perceived as opposed to two in the A1V2 condition, $t(7) = 2.09$, $p = .074$, $M = 421$ ms, $SE = 31$ ms for two-flash report; $M = 455$ ms, $SE = 28$ ms for one-flash report; Bonferroni adjusted α was $\alpha/2 = .025$. Mean RT as a function of perceived number of flashes for A2V1 and A1V2 trials is shown in Figure 3.

Next, we carried out the following comparisons: A1V2 when two flashes were perceived compared with the A1V1 condition, when one flash was reported in A2V1 trials compared with A2V2, and when two flashes versus one was perceived in the V2 trials; Bonferroni-adjusted α was $\alpha/3 = .0167$. First, a paired-samples *t* test revealed that when two flashes were perceived in A1V2 trials, mean RTs were not different in comparison with the A1V1 condition, $t(7) = 1.62$, $p = .15$. Second, the *t* test comparing A2V1 (one flash) with A2V2 was also nonsignificant, $t(7) < 1$, $p = .86$. Third, we carried out a paired-samples *t* test within the V2 condition comparing RTs when two flashes were reported versus when one was reported. Results failed to show evidence for significant results, $M = 456$ ms, $SE = 82$ ms, for two-flash report; $M = 438$ ms, $SE = 86$ ms for single-flash report, $t(7) < 1$, $p = .43$.

Individual Participant Analysis
To investigate these effects in greater detail, we conducted an analysis of individual participant data. These comparisons were carried out at the individual level to examine variability in information processing strategies, that is, whether multiple percepts or cross-modal facilitation contributes to changes in audiovisual processing speed in an observer. In doing so, we carried out statistical tests comparing entire RT distributions derived from the different trial types. Specifically, Kolmogorov–Smirnov (KS)

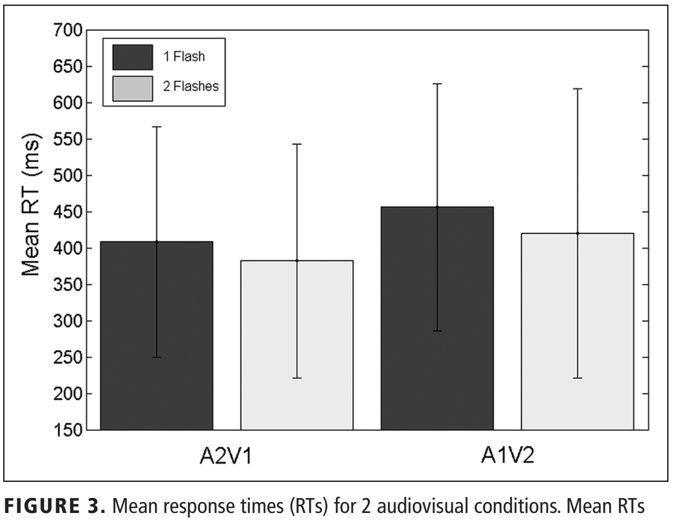

FIGURE 3. Mean response times (RTs) for 2 audiovisual conditions. Mean RTs were broken down by the number of reported flashes for the A2V1 and A1V2 conditions. Error bars indicate 90% confidence intervals

tests (see Kolmogorov, 1933) can be used to determine whether two experimental conditions or trial types differ significantly vis-à-vis their cumulative distribution functions.

RT AS A FUNCTION OF PERCEPTUAL REPORTS.

We implemented one-sided KS tests at the level of the RT distribution to determine whether each individual showed significant results for the following five effects: A2V1 (one vs. two flashes), A1V2 (one vs. two flashes), V2 (one vs. two flashes), A1V1 versus A1V2 (two flashes), and A2V1 (one flash) versus A2V2. The results for these tests are shown in Table 3. First, for the A2V1 (one vs. two flashes) comparison, half of the participants (4, 5, 7, and 8) showed evidence for significantly faster responses when two flashes were reported as opposed to only one. Recall that each of the eight observers showed evidence for the fission effect in terms of reported number of flashes. Second, for the A1V2 comparison (one vs. two flashes), half of the participants (1, 2, 6, and 8) yielded faster responses when two flashes were reported. Importantly, each participant who showed this RT effect also showed strong evidence for a significant fusion effect. Only one participant (i.e., 5) showed evidence for a fusion effect without a corresponding audiovisual RT effect. Results from these two tests are consistent with the hypothesis that RT effects were driven by the fusion or fission effect.

To examine whether a greater number of perceptual responses (one vs. two perceived flashes) may have driven results, we carried out three more statistical tests: The statistical results for V2 (one vs. two flashes) indicate that three participants responded faster when two flashes were perceived (5, 7, and 8). For the A1V1 versus A1V2 (two flashes) and A2V1 (one flash) versus A2V2 tests, Participant 7 showed significantly faster responses for A1V2 (two perceived flashes) trials, and Participants 2, 5, and 8 showed significantly faster results for the A2V2 trials, respectively. These results suggest that participants' response rates were at least susceptible to changes in processing speed as a function of perceived visual flashes. The following section summarizes the combined results from the individual analyses on perceived number of flashes, and statistical tests, in terms of our two theories.

TABLE 3. 1-Sided Kolmogorov–Smirnov Test Statistics and Corresponding *p* Values

Participant	A2V1, 1 vs. 2 (*p*)	A1V2, 1 vs. 2 (*p*)	V2, 1 vs. 2 (*p*)
1	.10 (.29)	.26 (.02)*	.09 (.88)
2	.08 (.42)	.30 (.02)*	.10 (.39)
3	.11 (.48)	.08 (.65)	.04 (.17)
4	.31 (<.001)*	.12 (.71)	.21 (.33)
5	.56 (.005)*	.17 (.30)	.47 (.01)*
6	.10 (.31)	.46 (.002)*	.16 (.12)
7	.16 (.03)*	.27 (.25)	.37 (.02)*
8	.45 (.002)*	.56 (.05)*	.41 (.001)*

Participant	A1V1 vs. A1V2, 2 flashes (*p*)	A2V1, 1 flash, vs. A2V2 (*p*)
1	.05 (.82)	.14 (.32)
2	.09 (.65)	.99 (<.001)*
3	.01 (.99)	.01 (.99)
4	.01 (.99)	.04 (.93)
5	.01 (.99)	.33 (.001)*
6	.36 (.02)	.16 (.09)
7	.02 (.93)*	.12 (.17)
8	.23 (.61)	.45 (.004)*

These tests were carried out to determine whether the right-hand quantity was significantly larger than the left-hand quantity.
**Statistically significant.*

Summary of Analyses: What Models Were Supported?

Table 4 provides a summary of the statistical test results. These indicate whether each participant was susceptible to the fission or fusion effects, whether they showed faster responses to two versus one flash for different trial types (the five statistical tests from the previous section), and given this information, whether evidence supports Theory 1, Theory 2, or both.

Participant 1 showed evidence for a significant fission and fusion effect while showing slower responses only for one perceived flash in the A1V2 condition. Because a significant effect in RTs was observed for A1V2 along with the fusion effect, Theory 2 provides the best overall explanation. This is because evidence for fusion exists without any independent evidence for faster responses as a function of increased perceptual judgments. Similar logic also explains the results

TABLE 4. Summary and Conclusions From Individual Participant Analyses

Participant	Fission	Fusion	A2V1	A1V2	A1V1 vs. A1V2	A2V1 vs. A2V2	V2	Theory 1?	Theory 2?
1	Yes	Yes	∅	Yes	∅	∅	∅	No	Yes
2	Yes	Yes	∅	Yes	∅	Yes	∅	Yes	Yes
3	Yes	∅	∅	∅	∅	∅	∅	No	No
4	Yes	∅	Yes	∅	∅	∅	∅	No	Yes
5	Yes	Yes	Yes	∅	∅	Yes	Yes	Yes	Yes
6	Yes	Yes	∅	Yes	∅	∅	∅	No	Yes
7	Yes	∅	Yes	∅	Yes	∅	Yes	Yes	Yes
8	Yes	Yes	Yes	Yes	∅	Yes	Yes	Yes	Yes

∅ = no evidence for this effect.

for Participants 4 and 6, who showed evidence for the fission illusion and for both the fission and fusion illusions, respectively.

Next, the configuration of results for Participants 2, 5, 7, and 8 is supported by both frameworks. This is because significant fission or fusion results were observed in conjunction with significantly shorter RTs in the A2V1 and A1V2 conditions (when two vs. one flash was perceived). Additionally, significant results were observed for at least one of the three tests examining the effects of the number of visual perceptual judgments. Together, these results indicate that the cross-modal connections responsible for the illusions or the number of perceptual judgments in the visual modality may drive processing speed results.

GENERAL DISCUSSION

This study investigated hypotheses about the origin of changes in redundant target processing speed by implementing an audiovisual detection study in the context of the sound-induced flash illusion. Using the sound-induced flash illusion to study multisensory binding is useful because it allows researchers to assess the effects of stimulus presentation versus perception on processing speed. We expanded the sound-induced flash illusion paradigm in order to examine whether the number of percepts, facilitatory auditory cueing, or both factors modulate audiovisual processing speed. This question has both theoretical and practical implications. For example, suppose auditory cues can significantly influence visual perception and processing rate or vice versa; this may suggest the efficacy of training auditory processing skills in visually impaired people and, likewise, training visual processing skills in listeners with hearing impairment.

One modification to previous research was the implementation of the full factorial speeded detection design to assess audiovisual processing speed. This was accomplished by investigating whether a processing time advantage occurred not only for one versus two perceived flashes in the fission A2V1 condition but also for one versus two perceived flashes in the fusion A1V2 condition. Participants also reported whether they saw one versus two flashes in the A1V1 and A2V2 conditions; however, two flashes and one flash were rarely reported in the A1V1 and A2V2 conditions, respectively.

In terms of perceived number of flashes, we replicated previous results showing that several participants reported two flashes when two auditory beeps and only one flash were presented. These results find converging support from neurophysiological data showing similar activation levels in the visual cortex in response to illusory and physically present double flashes (Watkins, Shams, Tanaka, Haynes, & Rees, 2006; Watkins, Shams, Josephs, & Rees, 2007). On the other hand, evidence for a fusion effect, in which the presentation of a single auditory stimulus in conjunction with two flashes has been reported to drive the perception of a single flash, was also observed in some participants. The previous literature on the fusion effect is mixed, and results across participants was variable in our study as well: Although several studies have successfully induced the fusion effect (Andersen et al., 2004, 2005; Shams et al., 2005; Watkins et al., 2007), some studies have failed to observe a significant effect for audiovisual fusions (e.g., Innes-Brown & Crewther, 2009; see also Bolognini, Convento, Fusaro, & Vallar, 2013, for a transcranial magnetic stimulation study). These results suggest the possibility that the fission and fusion illusions and processing speed related to perceptual effects are driven by different neurocognitive mechanisms. Critically, differences in factors such as stimulus duration and temporal structure (e.g., stimulus onset asynchrony) may contribute to the divergent findings between studies (see Apthorp, Alais, & Boenke, 2013).

Next, the processing time results showed that overall, processing speed was faster when two flashes were perceived as opposed to only one. For the individual participant analyses, results indicated that several participants showed faster responses when two flashes were perceived for A2V1 trials or for A1V2 trials. Crucially, the pattern of results for individual observers lined up with the predictions provided by the auditory cueing theory (Theory 2) or both theories. This suggests that although the number of percepts may contribute to multisensory processing speed in certain cases, the most common facilitator is cross-modal connections (e.g., Eidels et al., 2011) that increase processing speed and also probably contribute to the perception of the double flash.

Conclusion

Our data support the cross-modal facilitation theory, as seven of eight participants showed effects consistent with the hypotheses (e.g., Eidels et al., 2011). Importantly, evidence also suggests that at least some observers may also benefit from a combination of cross-modal facilitation in addition to two rather than only one visual percept (refer to Theory 1). Critically, when the sound-induced flash illusion did occur, evidence was obtained for faster responses when two versus one flash was present across several participants. Only one participant, 3, failed to show significant effects for any of the processing speed tests. Nonetheless, results indicate that the majority of observers should benefit from cross-modal cueing, multiple percepts, or a combination of both factors. Finally, we did not obtain evidence showing that multiple percepts alone contribute to faster multisensory processing speed.

NOTES

This project was supported by an internal University Research Office grant awarded to the first author at Idaho State University, and the IDeA Network of Biomedical Research Excellence Program, National Institutes of Health grants P20 RR016454 (National Center for Research Resources) and P20 GM103408 (National Institute of General Medical Sciences).

Address correspondence about this article to Nicholas Altieri, Idaho State University, 921 S. 8th Avenue, Stop 8116, Pocatello, ID 83209 (e-mail: altinich@isu.edu).

REFERENCES

Altieri, N., & Townsend, J. T. (2011). An assessment of behavioral dynamic information processing measures in audiovisual speech perception. *Frontiers in Psychology, 2*(238), 1–15.

Altieri, N., & Wenger, M. (2013). Neural dynamics of audiovisual integration efficiency under variable listening conditions: An individual participant analysis. *Frontiers in Psychology, 4*(615), 1–15.

Andersen, T. S., Tiippana, K., & Sams, M. (2004). Factors influencing audiovisual fission and fusion illusions. *Cognitive Brain Research, 21,* 301–308.

Andersen, T. S., Tiippana, K., & Sams, M. (2005). Maximum likelihood integration of rapid flashes and beeps. *Neuroscience Letters, 380,* 301–308.

Apthorp, D., Alais, D., & Boenke, L. T. (2013). Flash illusions induced by visual, auditory, and audiovisual stimuli. *Journal of Vision, 13*(5), 1–15.

Bolognini, N., Convento, S., Fusaro, M., & Vallar, G. (2013). The sound-induced phosphine illusion. *Experimental Brain Research, 231,* 469–478.

Eidels, A., Houpt, J., Altieri, N., Pei, L., & Townsend, J. T. (2011). Nice guys finish fast and bad guys finish last: A theory of interactive parallel processing. *Journal of Mathematical Psychology, 55,* 176–190.

Elze, T. (2010). Misspecifications of stimulus presentation durations in experimental psychology: A systematic review of the psychophysics literature. *PLOS ONE, 5*(9), e12792.

Fiedler, A., O'Sullivan, J., Schroter, H., Miller, J., & Ulrich, R. (2011). Illusory double flashes can speed up responses like physical ones: Evidence from the sound-induced flash illusion. *Experimental Brain Research, 214,* 113–119.

Giray, M., & Ulrich, R. (1993). Motor coactivation revealed by response force in divided and focused attention. *Jour-*

nal of Experimental Psychology: Human Perception and Performance, 19, 1278–1291.
Grice, G. R., Canham, L., & Gwynne, J. W. (1984). Absence of a redundant-signals effect in a reaction time task with divided attention. *Perception & Psychophysics, 36,* 565–570.
Innes-Brown, H., & Crewther, D. (2009). The impact of spatial incongruence on an auditory–visual illusion. *PLOS ONE, 4*(7), e6450.
Kolmogorov, A. (1933). Sulla determinazione empirica di una legge di distribuzione [On the empirical determination of a distribution law]. *Giornale dell'Istituto Italiano degli Attuari, 4,* 83–91.
Miller, J. O. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology, 14,* 247–279.
Miller, J. O. (1986). Timecourse of coactivation in bimodal divided attention. *Perception & Psychophysics, 40,* 331–343.
Mishra, I., Martinez, A., & Hillyard, S. A. (2008). Cortical processes underlying sound-induced flash fusion. *Brain Research, 1242,* 102–115.
Mordkoff, J. T., & Yantis, S. (1991). An interactive race model of divided attention. *Journal of Experimental Psychology: Human Perception and Performance, 17,* 520–538.
Otto, T. U., & Mamassian, P. (2012). Noise correlations in parallel perceptual decision making. *Current Biology, 22,* 1391–1396.
Raab, D. H. (1962). Statistical facilitation of simple reaction times. *Transactions of the New York Academy of Sciences, 24,* 574–590.
Savazzi, S., & Marzi, C. A. (2008). Does the redundant signal effect occur at an early visual stage? *Experimental Brain Research, 184,* 275–281.
Schroter, H., Fiedler, A., Miller, J. O., & Ulrich, R. (2011). Fusion prevents the redundant signals effect: Evidence from stereoscopically presented stimuli. *Journal of Experimental Psychology: Human Perception and Performance, 37,* 1361–1368.
Schroter, H., Frei, L. S., Ulrich, R., & Miller, J. (2009). The auditory redundant signals effect: An influence of number of stimuli or number of percepts? *Attention, Perception, & Psychophysics, 71,* 1375–1384.
Shams, L., Kamitani, Y., & Shimojo, S. (2000). What you see is what you hear. *Nature, 408,* 788.
Shams, L., Ma, W. J., & Beierholm, U. (2005). Sound-induced flash illusion as an optimal percept. *NeuroReport, 16,* 1923–1927.
Townsend, J. T., & Altieri, N. (2012). An accuracy–response time capacity assessment function that measures performance against standard parallel predictions. *Psychological Review, 119,* 500–516.
Townsend, J. T., & Nozawa, G. (1995). Spatio-temporal properties of elementary perception: An investigation of parallel, serial and coactive theories. *Journal of Mathematical Psychology, 39,* 321–360.
Townsend, J. T., & Wenger, M. J. (2004). A theory of interactive parallel processing: New capacity measures and predictions for a response time inequality series. *Psychological Review, 111,* 1003–1035.
Ulrich, R., Miller, J., & Schroter, H. (2007). Testing the race model inequality: An algorithm and computer programs. *Behavior Research Methods, 39,* 291–302.
Watkins, S., Shams, L., Josephs, O., & Rees, G. (2007). Activity in human V1 follows multisensory perception. *NeuroImage, 37,* 572–578.
Watkins, S., Shams, L., Tanaka, S., Haynes, J. D., & Rees, G. (2006). Sound alters activity in human V1 in association with illusory visual perception. *NeuroImage, 31,* 1247–1256.