

## Auditory enhancement of visual temporal order judgment

By: W. David Hairston, Donald A.Hodges, Jonathan H. Burdette, and Mark T.Wallace

Hairston, W., [Hodges, D.](#), Burdette, J. & Wallace, M. (2006) Auditory enhancement of visual temporal order judgment. *NeuroReport*. May 29; 17:8, 791-5.

Made available courtesy of LIPPINCOTT WILLIAMS & WILKINS:

<http://journals.lww.com/neuroreport/pages/default.aspx>

**This format of the article is not the final published version.**

**\*\*\*Note: Figures may be missing from this format of the document**

### **Abstract:**

Although numerous studies have shown that response times can be speeded by the presentation of multisensory stimuli, here we show that such speeding can be seen even when the second sensory channel fails to provide any task-relevant (i.e. redundant) information, and where cueing appears an unlikely explanation. Study participants performed a visual temporal order judgment task in the presence of task uninformative auditory cues, with the latter sound delayed relative to the latter visual cue. Responses were maximally speeded when the auditory stimulus was delayed by a short time (i.e. 100 ms) relative to the second visual target. These results illustrate a unique form of temporal benefit underlying a multisensory interaction, and form the basis for a novel explanation of these perceptual enhancements. Keywords: cross-modal; multisensory; response time; temporal enhancement

### **Article:**

#### *Introduction*

It is well documented that combining stimuli from different senses can have dramatic effects on human perception and performance. Some of these benefits include improvements in target detectability [1,2], perceived stimulus intensity [3,4], and target localization [5,6], as well as a marked speeding of responses [7,8]. Typically, these perceptual enhancements have been found to depend upon the stimuli being within close spatial and temporal proximity [9–13], although interactions have been observed in some cases even with significant spatial disparities [3,14,15].

In another illustration of the perceptual benefits of multisensory cues, Morein-Zamir and colleagues [16] have recently shown that auditory cues can improve an individuals' ability to discriminate the temporal order of two spatially distinct visual stimuli. They found that the introduction of a slight delay between the latter of the two visual targets and the second of two non-relevant auditory cues resulted in improved accuracy when compared with conditions in which there was no delay, despite the uninformative nature of the auditory cues. The effect occurred only when the delay was very short (100 ms), and has been interpreted as an example of 'temporal ventriloquism' [16], whereby the second auditory cue delays the perceptual appearance of the latter visual target.

Along with predicting improvements in response accuracy, temporal ventriloquism also predicts a slowing of response times (RTs) that directly reflects the perceptual delay in the appearance of the second visual target. The current study was designed to test this prediction, and also provided a unique opportunity to examine multisensory improvements in RTs under conditions in which information from one of the sensory channels is uninformative for the task.

#### *Methods*

##### **Study participants**

Twenty individuals (five females; mean age 33.8 years) participated; all were paid and provided written informed consent. Procedures were approved by the Wake Forest University Internal Review Board, and conformed to the 1964 Declaration of Helsinki.

### **Task overview**

The task was similar to that previously described [17]. Participants performed a visual temporal order judgment (TOJ) task in which they were asked to report which of two circles appeared first. The experimental procedure was broken into three blocks. The first block was a staircase procedure to obtain a threshold stimulus onset asynchrony (SOA) for visual TOJ performance. The second block was a verification of this derived threshold, and the third block consisted of test trials, many of which included task-irrelevant auditory signals.

### **Stimuli**

Visual stimuli consisted of two white circles (each covering approximately 3° of visual angle) against a black background presented on a computer monitor (LG 915FT+, 200 Hz vertical scan). The circles appeared 10° above and 10° below a continually illuminated fixation cross. Auditory stimuli (broadband noise, 20 Hz–20 kHz, 65 dB sound pressure level, 10 ms duration) were presented binaurally via headphones, and were simultaneous and of equal amplitude. Stimuli were presented using E-Prime presentation software (Psychology Software Tools Inc., Pittsburgh, Pennsylvania, USA). The temporal consistency of all stimuli was externally verified before data acquisition.

#### **Block 1 – staircase procedure: visual temporal order judgment**

Participants were given instructions and allowed a brief practice session. During each trial, participants maintained fixation and after a delay of 1500 ms, the first circle appeared. Following a variable SOA, the second circle appeared at the other location. The participant's task was to respond, using a button press, as to which of the circles appeared first. Following the response, both circles disappeared, and a new trial began. For all blocks, participants were asked to be as accurate as possible, but also to respond quickly. It was recommended that if they were uncertain, they were to take their best guess; they were not given any feedback.

An adaptive staircase procedure was used to determine the SOA necessary to perform the visual TOJ task at threshold. Three staircases ran independently, starting at SOAs of 15, 55, and 85 ms. For each, the initial step size was 20 ms and decreased to 10 ms after five reversals, and then to 5 ms after the next four reversals. The SOA increased one step after each incorrect response, and decreased one step after two consecutive correct responses. Each staircase terminated after 16 reversals, and an average was calculated from the last five reversal values. An overall average was then determined for the three staircases and rounded up to the nearest value compatible with the vertical scan rate of the monitor (i.e. 5 ms increments). This process converged on perceptual thresholds with average performance rates of ~70–75% accuracy.

#### **Block 2 – staircase verification: visual temporal order judgment**

In order to verify that the SOA derived from the staircase procedure was near threshold for each individual, participants performed a second block, with SOA values set at 10 ms above and 10 ms below their previously determined thresholds. Each of these three SOAs was repeated 20 times in random order. This sequence was repeated with new values if performance was not near threshold (i.e. between 68 and 78%).

#### **Block 3 – experimental procedure: visual temporal order judgment with auditory cues**

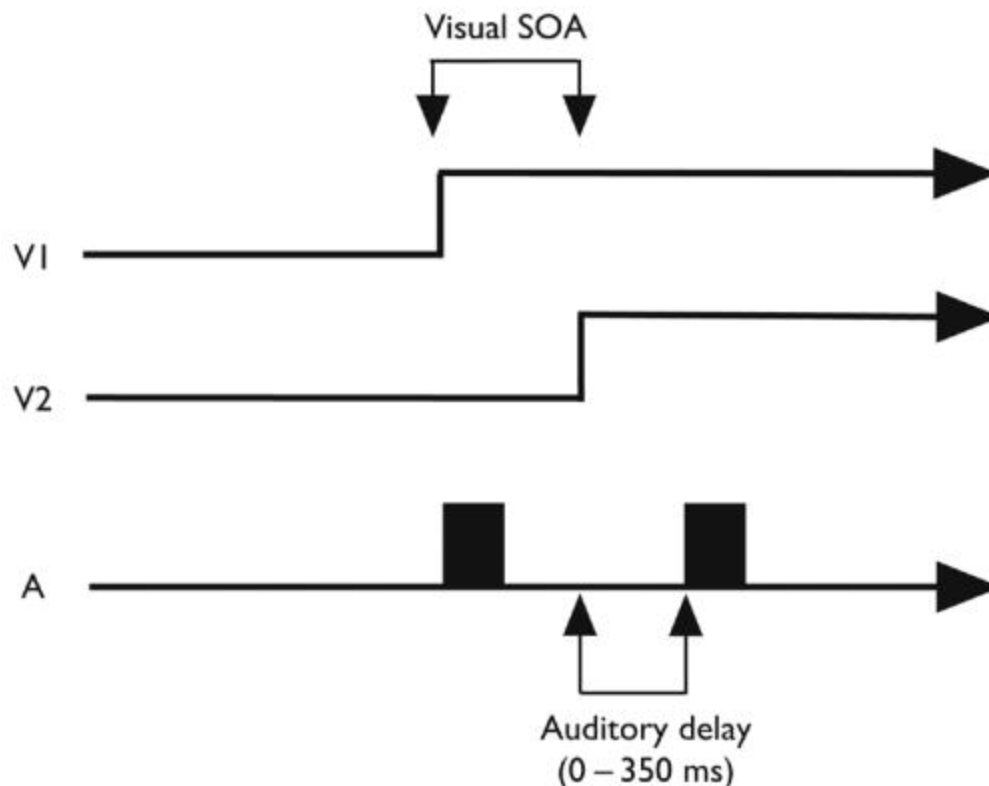


Fig. 1 Schematic of the trial sequence. Two visual stimuli (V1 and V2, top and middle lines) appear on a video monitor separated by a short stimulus onset asynchrony (SOA). Participants are asked to report which appeared first. During experimental testing, the SOA is fixed at that individual's threshold, and additional auditory stimuli (A, bottom line) are presented with a variable delay (0–350 ms) between the onset of the second visual target and second auditory stimulus.

Figure 1 illustrates the trial sequence. On each trial, the SOA between the two visual stimuli was fixed according to the staircase-derived threshold. Additionally, two identical sounds were presented through headphones on most trials. While the onset of the first sound was always synchronous with the onset of the first visual target, the onset of the second sound was delayed by 0–350 ms (50 ms steps) relative to the onset of the second visual target. A randomly interleaved no-sound condition provided baseline (i.e. visual only) performance. Each auditory delay, plus the no-sound condition, was presented 36 times in pseudorandom order, with the presentations split between two 5 to 6-min blocks. Participants were instructed to perform the same task as they had previously; however, they were told that they would also hear sounds. As the sounds were presented binaurally through headphones with no interaural timing or level differences, they did not provide any spatial information useful for the task. The auditory cues, however, did provide temporal information.

### Data analysis

Both response accuracy and RT were recorded for each trial, with RTs calculated relative to the onset of the second visual cue. Only data from the final procedure (visual TOJ with auditory cues) were compiled for group analyses. All trials with RTs >3 s were removed. For each participant, the median RT was computed for each of the eight experimental conditions, and group averages were calculated. All experimental comparisons were carried out using paired t-tests with an [alpha] criterion of  $P < 0.05$ . Initially, all responses were included, but during later analyses (see Results) this was restricted to include only correct responses.

### Results

#### Response accuracy

Similar to previous reports [17], when participants were presented with spatially non-informative auditory stimuli whose onsets were simultaneous with the onset of the visual targets, response accuracy was not significantly different when compared with trials in which only the visual targets were presented, although a

trend toward improvement was apparent [ $t(19)=1.90$ ,  $P=0.072$ ]. In addition, similar to previous reports [16,17], when a small delay was introduced between the onset of the second visual target and the onset of the second auditory stimulus, accuracy improved. The largest decrease in RT is observed when the visual-auditory delay is 100 ms [ $t(19)=5.93$ ,  $P<0.05$ ].

### Response time

Median RTs were computed for each participant and condition, and group means are shown in Fig. 2. Participants showed a consistent reduction in median RT for simultaneous visual-auditory trials (i.e. 0 ms delay) compared with the visual-alone condition [ $t(19)=4.1$ ,  $P<0.05$ ]. Additionally, introducing a small delay between the second visual target and the second auditory stimulus further speeded RTs, but only within a specific temporal 'window' (Fig. 2a, asterisk). A significant decrease in RT is observed when the visual-auditory delay is 100 ms [ $t(19)=5.93$ ,  $P<0.05$ ], the delay at which the largest effect is seen. As the delay is further increased, RTs increase steadily and are statistically indistinguishable from baseline by 300 ms [ $t(19)=0.59$ ,  $P>0.05$ ].

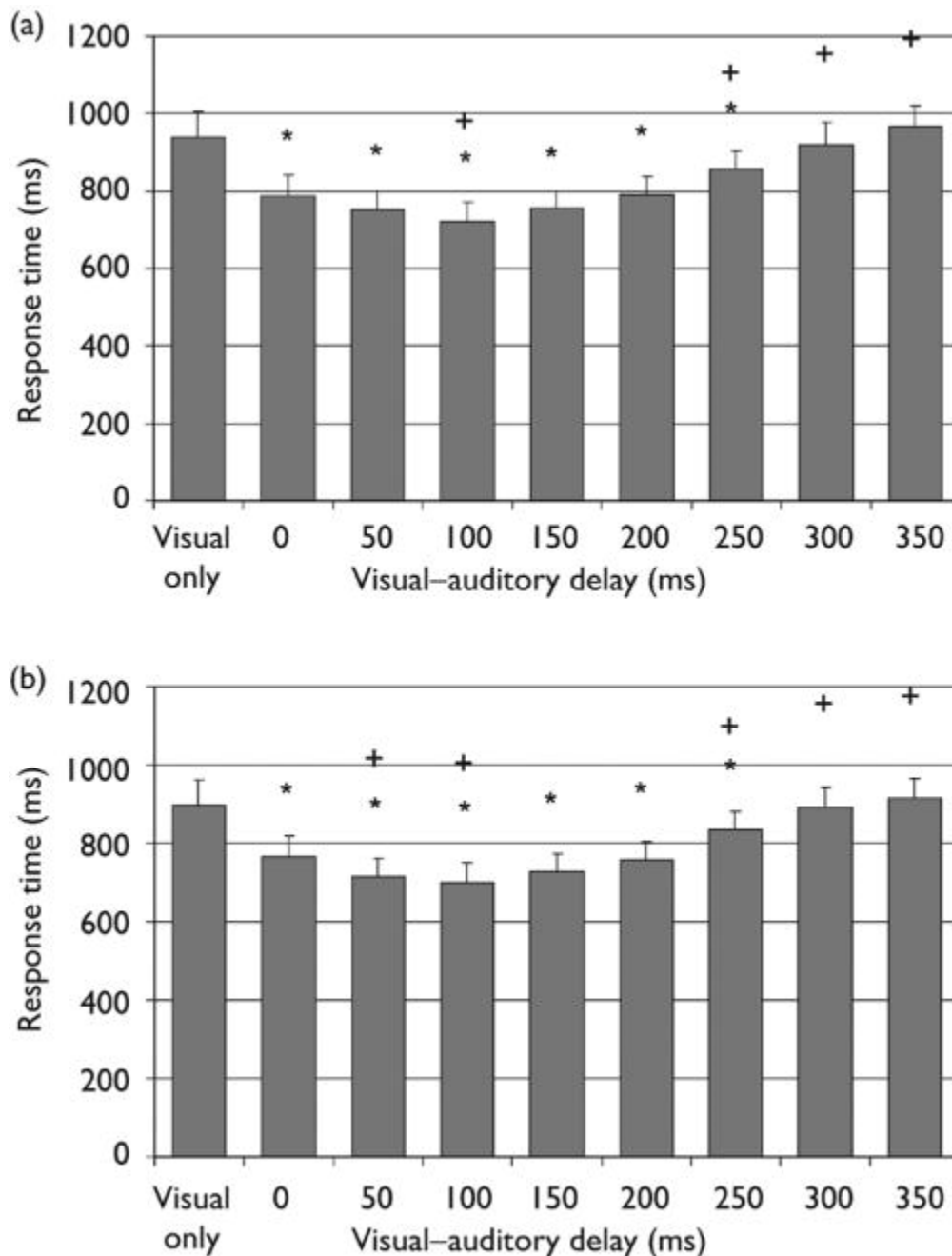


Fig. 2 (a) Average response times (RT) as a function of experimental condition (i.e. visual–auditory delay). Included are all responses (i.e. correct and incorrect trials). Asterisks (\*) denote conditions in which RTs are significantly ( $P < 0.05$ ) faster than for the visual stimulus alone (i.e. visual only). Plus signs (+) denote conditions in which RTs are significantly different from the simultaneous (i.e. 0 ms delay) multisensory condition. Note that the fastest responses are seen when there is a small visual–auditory delay (i.e. 100 ms). Error bars represent the group SEM. (b) Same analysis and conventions in (a) except that all incorrect response have been removed. Note the very similar patterns in (a) and (b).

One potential criticism of these comparisons is that they include both differences attributable to the different modalities of stimulus delivery (visual vs. multisensory) and differences in the temporal structure of the stimulus complex. Consequently, as an additional comparison, conditions in which RTs were significantly different from the synchronous (0 ms) visual–auditory condition are also illustrated in Fig. 2a (plus). Such a comparison highlights differences attributable only to the temporal structure of the stimuli. Using this alternative baseline, not only is there a significant speeding of RTs with the 100 ms visual–auditory delays, but there is also a significant slowing of RTs with longer (i.e. 250–350 ms) delays.

### **Response times and accuracy**

As the speed with which participants respond on a given task can often be closely related to their ability to perform the task, we wanted to ensure that the effect on RTs was not a secondary effect related to participant's changes in accuracy. Simply stated, participants could be slower to respond when they are less sure of the correct response. To examine this potential confound, RT data were re-analyzed with incorrect responses removed (Fig. 2b).

A noticeable decline was observed in RT when error trials were removed [ $F(1,19)=70.96$ ,  $P < 0.05$ ], illustrating that participants are indeed faster to respond when they know the correct response. Nonetheless, the pattern of changes in RT as a function of visual–auditory delay was similar for both data sets (compare Fig. 2a and b). As a final step in examining the potential interaction between changes in accuracy and RT, we used stepwise linear regression to examine the relationship between RT and visual–auditory delay, with the semi-partial correlation to response accuracy accounted for within the model. Using this method, visual–auditory delay remained a reliable, significant predictor of participants' RT [ $t(157)=4.67$ ,  $P < 0.05$ ], even with overlapping variance associated with response accuracy removed. This led to a significant semi-partial correlation between delay and RT, ( $r=0.33$ ,  $P < 0.05$ ), suggesting a modest degree of independence between RT and accuracy.

### **Response time distributions**

As an additional evaluation of the influence of auditory cues on visual performance, we examined the average change in the shape of the RT distribution for each cross-modal condition by calculating the temporal interval between the 25th and 75th quartiles for each participant's RTs. We found a consistent decrease in the breadth of the response distribution that paralleled the changes in median RT. Figure 3a shows these distributions for a representative participant (0 and 100 ms delay conditions), while Fig 3b shows the group averages for all conditions. Note that the smallest average interquartile range is seen at the 100 ms delay, a significant decline from visual-only [ $t(19)=2.92$ ,  $P < 0.05$ ] or simultaneous [ $t(19)=2.27$ ,  $P < 0.05$ ] multisensory conditions.

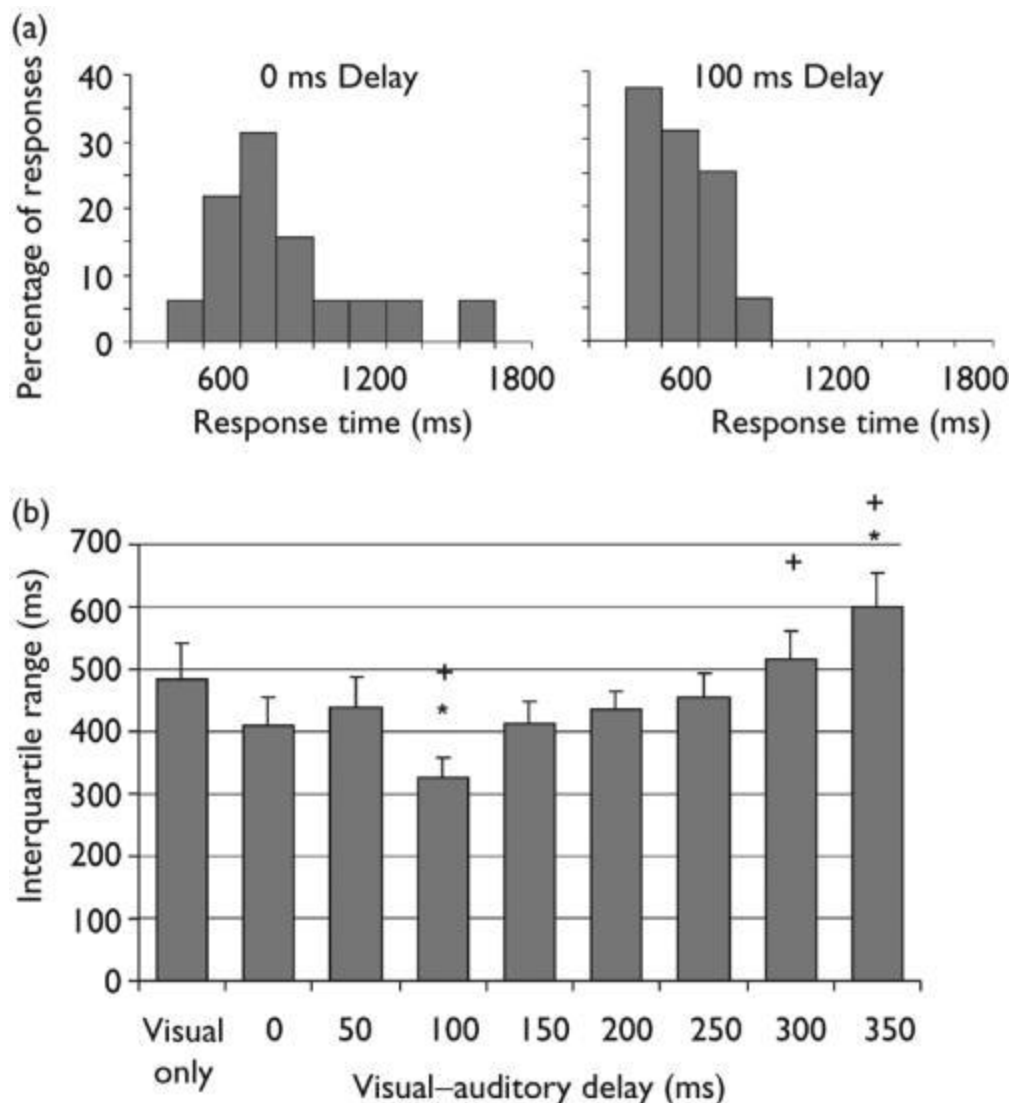


Fig. 3 (a) Distribution of response times for a representative participant for the 0 ms (left) and 100 ms (right) delay conditions. Note the substantial narrowing of the distribution for the 100 ms delay condition. (b) Average group data depicting the change in the breadth of the response distributions (i.e. interquartile range) for each of the experimental conditions. Note again the narrowing of these distributions at short visual-auditory delays. Symbols represent the same statistical comparisons as in Fig. 2.

### Discussion

Here, we show that the addition of a task-irrelevant auditory stimulus to a visual temporal-processing task can result in a substantial speeding of responses. Responses were maximally speeded when there was a small delay between the second visual target and the second auditory stimulus, a pattern very similar to what has been previously reported for improvements in response accuracy using a similar paradigm [16,17]. Intriguingly, these perceptual improvements are seen despite the fact that the auditory stimulus is not directly informative for the task at hand.

The current paradigm provided a unique opportunity to assess multisensory-mediated benefits in RTs in a context outside of traditional race model comparisons [10,18,19], as the second channel of sensory information (i.e. auditory) does not provide redundant information. Despite this, the auditory information clearly altered the speed at which participants performed the visual TOJ task, illustrating that the temporal relationship of auditory cues to visual targets can markedly influence not only the accuracy but also the speed of visual judgments. Consistent with the idea that the temporal structure of the stimuli is a key feature of the effect was the finding that performance was altered only for a certain set of temporal delays; a range that falls within the temporal

window in which other cross-modal neural, behavioral and perceptual effects have been reported [8,9,11,13,16,20,21].

Previous reports have attributed the accuracy improvements seen in this task to a form of ‘temporal ventriloquism’, in which the perceived occurrence of the second visual stimulus is shifted slightly in time by the presence of the second auditory stimulus [16]. While such an account is attractive, it seems incompatible with the speeded RTs found here. Rather, the expectation would be that responses should be somewhat slowed under circumstances in which accuracy is improved as a result of the delay in the perceived time of the onset of the second visual stimulus.

Consequently, we propose a somewhat different explanation, whereby the integration of visual and auditory information leads to enhanced visual temporal ‘acuity’. It is well established that perceptual temporal acuity in the visual system is relatively poor, particularly when compared with the auditory system, which is specialized for performing rapid temporal analyses [22,23]. Thus, we suggest that the temporal profile of a visual signal can be improved when integrated with an auditory cue that occurs in close temporal proximity, resulting in a speeding of RTs. Such an interpretation is directly supported by the narrowing of the temporal distributions seen when responses are speeded (i.e. Fig. 3).

One potential concern is that the speeded responses may simply be a reflection of a participant's improved performance accuracy; that is, perhaps they are more confident of their judgments and consequently respond faster. Arguing against this possibility is the finding that changes in RTs appear to be independent of participants' ability to perform the task, as the removal of incorrect trials showed an identical trend. Finally, by using a regression model, we have illustrated a degree of independence between response accuracy and RT. Intriguingly, this result suggests that these two performance measures may be mediated, at least in part, by distinct cross-modal interactive processes, despite the fact that they are ultimately bound into the same perceptual gestalt.

### *Conclusion*

The addition of a task-irrelevant but temporally informative auditory cue can dramatically affect not only a person's accuracy in performing a visual discrimination but also the speed with which they perform the task. We propose that this speeding occurs as a result of a temporal interaction between the representations of the visual and auditory stimuli that results in a sharpening of the temporal profile of the visual target.

### *References*

1. Lovelace CT, Stein BE, Wallace MT. An irrelevant light enhances auditory detection in humans: a psychophysical analysis of multisensory integration in stimulus detection. *Brain Res Cogn Brain Res* 2003; 17:447–453.
2. Bolognini N, Frassinetti F, Serino A, Ladavas E. ‘Acoustical vision’ of below threshold stimuli: interaction among spatially converging audiovisual inputs. *Exp Brain Res* 2005; 160:273–282.
3. Stein B, London N, Wilkenson L, Price D. Enhancement of perceived visual intensity by auditory stimuli: a psychophysical analysis. *J Cogn Neurosci* 1996; 8:497–506.
4. Odgaard EC, Arieh Y, Marks LE. Brighter noise: sensory enhancement of perceived loudness by concurrent visual stimulation. *Cogn Affect Behav Neurosci* 2004; 4:127–132.
5. Alais D, Burr D. The ventriloquist effect results from near-optimal bimodal integration. *Curr Biol* 2004; 14:257–262.
6. Hairston WD, Laurienti PJ, Mishra G, Burdette JH, Wallace MT. Multisensory enhancement of localization under conditions of induced myopia. *Exp Brain Res* 2003; 152:404–408.

7. Corneil BD, Van Wanrooij M, Munoz DP, Van Opstal AJ. Auditory-visual interactions subserving goal-directed saccades in a complex scene. *J Neurophysiol* 2002; 88:438–454. [Bibliographic Links](#) [[Context Link](#)]
8. Corneil BD, Munoz DP. The influence of auditory and visual distractors on human orienting gaze shifts. *J Neurosci* 1996; 16:8193–8207. [Bibliographic Links](#) [[Context Link](#)]
9. Lewald J, Guski R. Cross-modal perceptual integration of spatially and temporally disparate auditory and visual stimuli. *Brain Res Cogn Brain Res* 2003; 16:468–478.
10. Harrington LK, Peck CK. Spatial disparity affects visual-auditory interactions in human sensorimotor processing. *Exp Brain Res* 1998; 122:247–252.
11. Frens MA, Van Opstal AJ, Van der Willigen RF. Spatial and temporal factors determine auditory-visual interactions in human saccadic eye movements. *Percept Psychophys* 1995; 57:802–816.
12. Diederich A, Colonius H, Bockhorst D, Tabeling S. Visual-tactile spatial interaction in saccade generation. *Exp Brain Res* 2003; 148:328–337.
13. Slutsky DA, Recanzone GH. Temporal and spatial dependency of the ventriloquism effect. *Neuroreport* 2001; 12:7–10.
14. Hairston WD, Wallace MT, Vaughan JW, Stein BE, Norris JL, Schirillo JA. Visual localization ability influences cross-modal bias. *J Cogn Neurosci* 2003; 15:20–29.
15. Murray MM, Molholm S, Michel CM, Heslenfeld DJ, Ritter W, Javitt DC, et al. Grabbing your ear: rapid auditory–somatosensory multisensory interactions in low-level sensory cortices are not constrained by stimulus alignment. *Cereb Cortex* 2005; 15:963–974.
16. Morein-Zamir S, Soto-Faraco S, Kingstone A. Auditory capture of vision: examining temporal ventriloquism. *Brain Res Cogn Brain Res* 2003; 17:154–163.
17. Hairston WD, Burdette JH, Flowers DL, Wood FB, Wallace MT. Altered temporal profile of visual-auditory multisensory interactions in dyslexia. *Exp Brain Res* 2005; 166:474–480. [Bibliographic Links](#)
18. Miller J. Divided attention: evidence for coactivation with redundant signals. *Cognit Psychol* 1982; 14:247–279.
19. Raab DH. Statistical facilitation of simple reaction times. *Trans NY Acad Sci* 1962; 24:574–590.
20. Hughes HC, Reuter-Lorenz PA, Nozawa G, Fendrich R. Visual-auditory interactions in sensorimotor processing: saccades versus manual responses. *J Exp Psychol Hum Percept Perform* 1994; 20:131–153.
21. Stein BE, Meredith MA. *The merging of the senses*. Cambridge, Massachusetts: MIT Press; 1993.
22. Hirsh IJ, Sherrick CE Jr. Perceived order in different sense modalities. *J Exp Psychol* 1961; 62:423–432.
23. Welch RB, Warren DH. Immediate perceptual response to intersensory discrepancy. *Psychol Bull* 1980; 88:638–667.