

**On the Spatial Specificity of Audiovisual  
Crossmodal Exogenous Cuing Effects**

Jae Lee & Charles Spence

*Crossmodal Research Laboratory, Department of  
Experimental Psychology, University of Oxford, Oxford, UK*

DATE: March, 2017

CORRESPONDENCE TO: Jae Lee, Crossmodal Research Laboratory, Department of  
Experimental Psychology, University of Oxford, Oxford, OX1 3UD, UK; E-MAIL:  
[jae.lee3@psy.ox.ac.uk](mailto:jae.lee3@psy.ox.ac.uk).

## ABSTRACT

It is generally-accepted that the presentation of an auditory cue will direct an observer's spatial attention to the region of space from where it originates and therefore facilitate responses to visual targets presented there rather than from a different position within the cued hemifield. However, to date, there has been surprisingly limited evidence published in support of such within-hemifield crossmodal exogenous spatial cuing effects. Here, we report two experiments designed to investigate within- and between-hemifield spatial cuing effects in the case of audiovisual exogenous covert orienting. Auditory cues were presented from one of four frontal loudspeakers (two on either side of central fixation). There were eight possible visual target locations (one above and another below each of the loudspeakers). The auditory cues were evenly separated laterally by  $30^\circ$  in Experiment 1, and by  $10^\circ$  in Experiment 2. The potential cue and target locations were separated vertically by approximately  $19^\circ$  in Experiment 1, and by  $4^\circ$  in Experiment 2. On each trial, the participants made a speeded elevation (i.e., up vs. down) discrimination response to the visual target following the presentation of a spatially-nonpredictive auditory cue. Within-hemifield spatial cuing effects were observed only when the auditory cues were presented from the inner locations. Between-hemifield spatial cuing effects were observed in both experiments. Taken together, these results demonstrate that crossmodal exogenous shifts of spatial attention depend on the eccentricity of both the cue and target in a way that has not been made explicit by previous research.

*Keywords:* attention, spatial cuing, crossmodal, eccentricity, exogenous

## Introduction

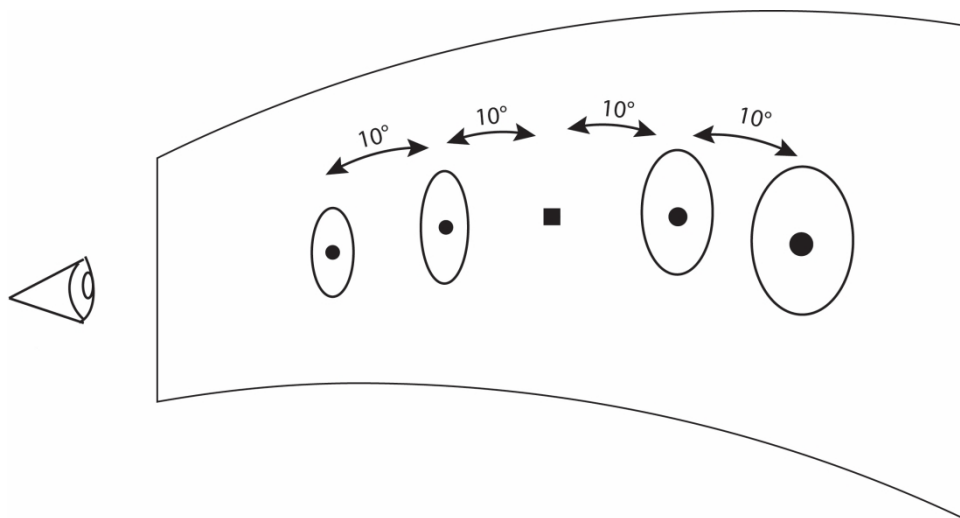
One of the most oft-replicated findings in the field of exogenous crossmodal spatial attention research is that the presentation of a task-irrelevant spatial cue in one sensory modality can facilitate the perception of targets presented in another modality on the same (i.e., cued) rather than opposite (i.e., uncued) side (this is typically referred to as the between-hemifield spatial cuing effect). Spatial cuing effects usually last for less than 300ms and are typically demonstrated in terms of faster RTs to targets presented ipsilateral to the preceding cues than those when presented contralaterally (e.g., Spence & Driver, 1997; see Spence, McDonald, & Driver, 2004, for a review). To date, crossmodal between-hemifield cuing effects have been documented between all possible combinations of visual, auditory, and tactile stimuli (e.g., when auditory targets are preceded by visual cues and vice versa; see Spence & McDonald, 2004; Spence et al., 2004, for reviews). Although the topic of *between*-hemifield spatial cuing effects has been investigated in detail (this being the default manipulation in virtually all crossmodal cuing studies that have been published to date), the spatial specificity of cuing effects within the cued hemifield (i.e., the *within*-hemifield cuing effect) has not been studied anything like so thoroughly.

Nevertheless, it has been argued by many authors that exogenous crossmodal cuing effects are spatially-specific (see Driver & Spence, 1998; Schmitt, Postma, & De Haan, 2001; Spence et al., 2004). That is, the presentation of a spatially-nonpredictive cue in one modality will direct a participant's attention to a specific location (or region of space, depending on the paradigm used) rather than to the entire hemifield in which the cue happens to have been presented (i.e., the left or right side of space; Spence et al., 2004). For instance, Schmitt et al. investigated whether changing the cue-target distance would modulate the magnitude of crossmodal spatial cuing effects. More specifically, they assessed distance effects using four cue-target combinations: visual-visual, visual-auditory, auditory-visual, and auditory-auditory. There were a total of four cue-target distance conditions in the study: 0-distance if the cue and target were presented at the same location; 1-distance if the cue and target were presented next to each other; 2-distance if there was an unused cue location between the presented stimuli; and 3-distance if the cue and target were presented from the outer-left and outer-right positions.

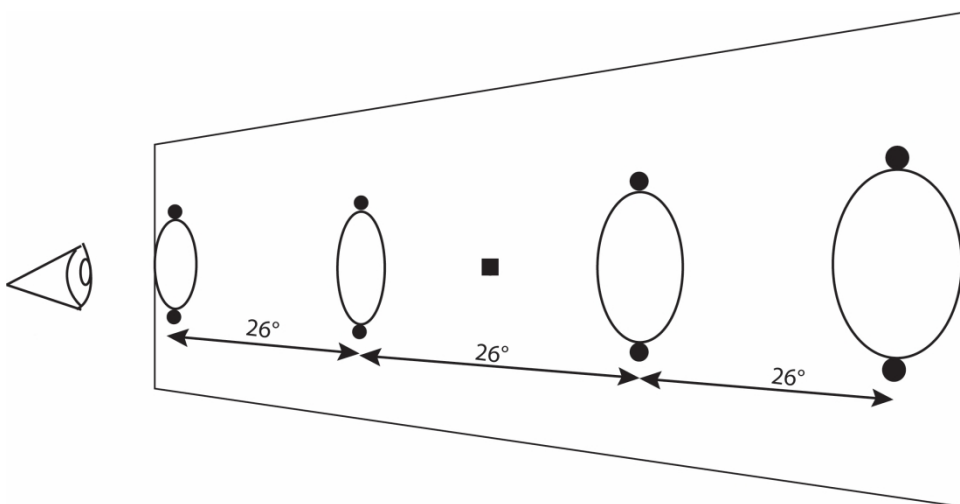
In Schmitt et al.'s (2001) study, an LED was placed at the centre of each of four loudspeakers. The stimuli were presented 10° (inner eccentricities) or 20° (outer eccentricities) to either

side of central fixation. All of the stimuli were presented at a fixed distance from the participants (i.e., on a circle centred on the participant's head; see Figure 1a). On each trial, a target was presented after a variable stimulus onset asynchrony (SOA; 125, 175, 300, 575, or 825ms), following the onset of the cue. The participants had to press one of the four laterally-placed response keys (e.g., an outer-left key for the outer-left target, an inner-right key for the inner-right target, etc.) to indicate the perceived location of the target as rapidly and accurately as possible.

(a)



(b)



*Figure 1.* (a) Schematic illustrations of Schmitt et al.'s (2001) experimental set-up and (b) Driver and Spence's (1998) orthogonal spatial-cuing paradigm. The four loudspeakers used to present the auditory cues are

shown as ellipses. Target LEDs are represented as black dots (a) in front of the loudspeakers or (b) above and below each loudspeaker. The fixation LED (shown here as a black square) was placed between the inner loudspeakers in both studies.

Schmitt et al. (2001) found that RTs were shortest when the visual targets were preceded by either visual or auditory cues from the same position (i.e., 0-distance condition) rather than from a different location (at all SOAs [125, 175, 300, 575, & 825ms]). Furthermore, RTs in the 1-distance condition were found to be faster than those in the 2-distance condition, which, in turn, were faster than those in the 3-distance condition. These results suggested the existence of a so-called distance effect. These findings add further support to the gradient model of spatial attention, according to which attentional facilitation is greatest at the cued location and decreases as a function of the distance from the cued location (e.g., Downing, 1988; Downing & Pinker, 1985; Mangun & Hillyard, 1988; Shulman, Wilson, & Sheehy, 1985, for visual gradients; Mock, Seay, Charney, Holmes, & Golob, 2015; Mondor & Zatorre, 1995; Teder-Sälejärvi & Hillyard, 1998, for auditory gradients; see also Kennett & Driver, 2014, for discussion in the context of visuotactile cuing). However, there are two possible explanations for the slower RTs in the 1- than the 0-distance condition. One is the within-hemifield spatial cuing effect; The other is that the distance effect between the 0- and the 1-distance might simply reflect a standard between-hemifield spatial cuing effect when auditory cues and visual targets just so happened to have been presented from the two inner locations (e.g., Spence & Driver, 1997). Therefore, in-and-of-themselves, distance effects such as these cannot be taken as evidence that the presentation of a spatially-nonpredictive auditory cue can shift visual attention in a spatially-specific manner within the cued hemifield.

Furthermore, it is important to note that the distance effects reported by Schmitt et al. (2001) were potentially confounded by response priming. Specifically, the dimension along which the auditory cues varied was the same as that of participants' responding. Some years ago, Spence and Driver (1994, 1997) pointed out that in certain non-orthogonal spatial discrimination tasks, faster RTs on cued trials than on uncued trials might simply reflect response priming rather than attentional facilitation. One popular method in the majority of spatial cuing studies that have been published to date in order to avoid response priming confound is the orthogonal spatial cuing paradigm (e.g., Ho, Tan, & Spence, 2006; Spence &

Driver, 1994, 1996, 1997). In those studies using this approach, the dimension along which the cue varies (e.g., left or right side) is intentionally made *orthogonal* to that of responses (e.g., discriminating upper vs. lower target locations). One key aim of the orthogonal cuing paradigm is to try and isolate any crossmodal attention facilitation from the possible confounding effects that might otherwise be attributable to response priming (see Spence & McDonald, 2004).

In fact, to date, no studies have investigated within-hemifield exogenous audiovisual spatial cuing effects without potential response bias concerns, except for a single study (Driver & Spence, 1998) described briefly in Spence et al. (2004) and Spence and McDonald (2004). In particular, these researchers reported an experiment in which the spatial specificity of audiovisual crossmodal cuing effects were assessed using the orthogonal cuing paradigm. In this case, the experimental set-up included two cue loudspeakers on each side with a pair of target LEDs, one placed above and the other below each loudspeaker (see Figure 1b). On each trial, the participants had to make a speeded elevation discrimination response (i.e., upper vs. lower) to the visual target presented from one of eight possible locations, while ignoring a spatially-nonpredictive auditory cue presented shortly beforehand. Spence et al. reported that “the presentation of an auditory cue [...] led to a spatially specific shift of attention that facilitated visual elevation discrimination response latencies maximally for visual targets presented from directly above and below the auditory cued location” (p. 9). Importantly, such within-hemifield spatial cuing effects were reported from all four of the possible cue positions (see Figure 2). RTs were 20-40ms faster when the cue and target were presented from the same position than from different lateral positions within the cued hemifield. Furthermore, RTs to visual targets tended to increase as a function of the cue-target distance, thus providing further support for the gradient model of attention (e.g., Downing & Pinker, 1985) and distance effects (e.g., Schmitt et al., 2001). It is, however, worth noting that these experimental findings have only ever been reported in review papers (e.g., Spence et al., 2004). As a result, the data supporting the spatial specificity of the audiovisual cuing effect has not been available for closer inspection.

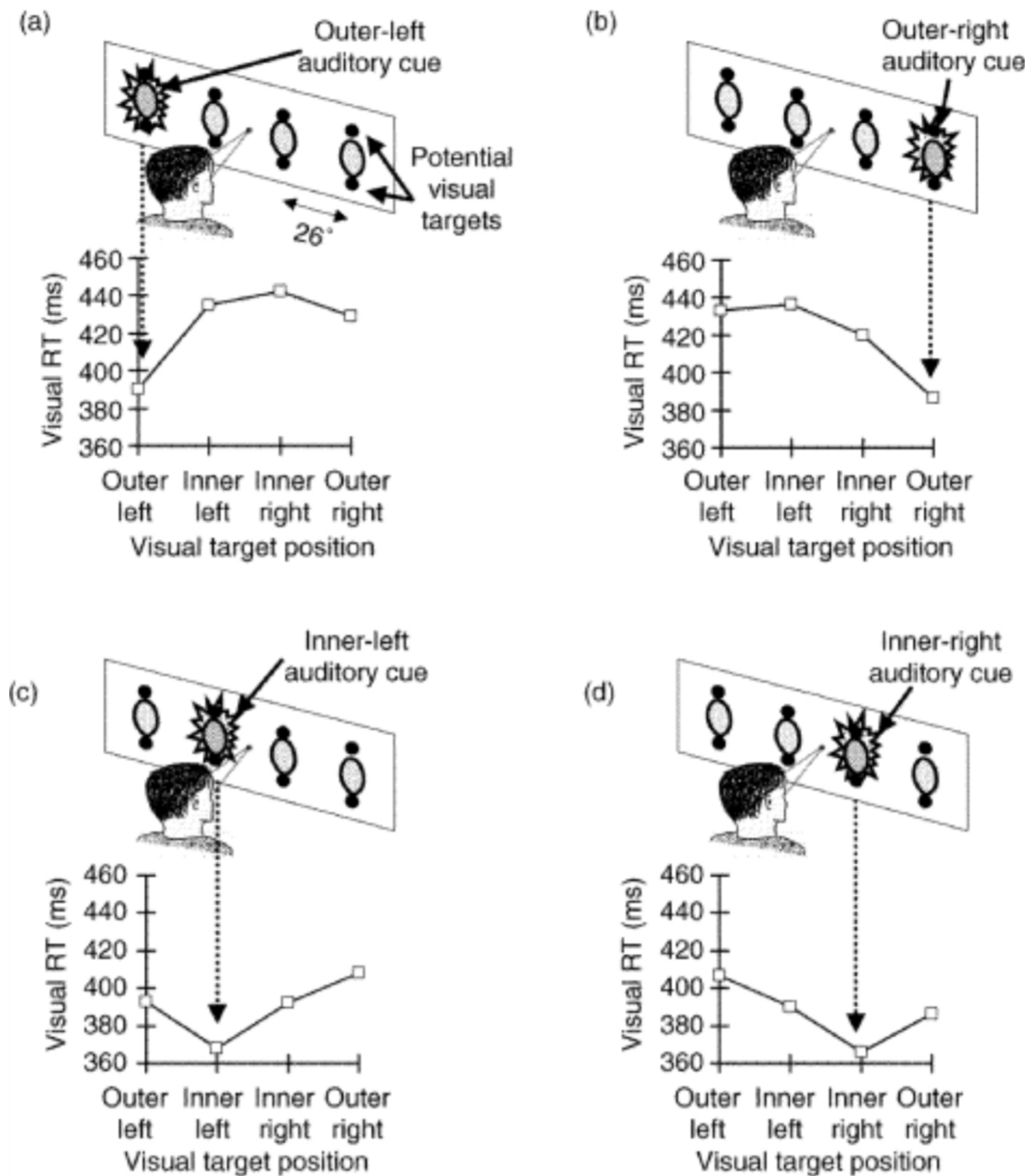


Figure 2. Mean reaction times (RTs) in milliseconds (ms) in the (a) outer-left, (b) outer-right, (c) inner-left, and (d) inner-right auditory cue conditions as a function of the visual target positions (reprinted with permission from Spence et al., 2004, p. 286).

In summary, the evidence that has been published to date indicates that localised auditory cues can pull attention to the cued hemifield, either exogenously (e.g., Spence & Driver, 1997)

or else endogenously (e.g., Spence & Driver, 1996). Furthermore, audiovisual spatial cuing effects are reported to be larger when the cue and target are presented from the same, rather than from different lateral positions within the cued hemifield (see Driver & Spence, 1998; also see Spence et al., 2004). However, to date, only limited evidence has been provided in support of the existence of within-hemifield audiovisual spatial cuing effects. Furthermore, to the best of our knowledge, there has been no audiotactile or visuotactile cuing study on within-hemifield spatial cuing effects (though see Kennett & Driver, 2014<sup>1</sup>).

Therefore, in order to better understand the spatial specificity of audiovisual exogenous cuing effects (which might be important in both an applied and theoretical context; e.g., see Baldwin, Spence, Bliss, Brill, Wogalter, Mayhorn, & Ferris, 2012; Ho & Spence, 2008, for reviews), two experiments were designed to investigate within-hemifield spatial cuing and between-hemifield spatial cuing. The experimental set-up was closely modelled on that reported almost two decades ago by Driver and Spence (1998; i.e., two possible auditory cue positions on each side, and two possible visual target locations; one directly above and another directly below each loudspeaker; see Figure 3). Because the spatial resolution in vision is known to be better at the fovea than in the periphery (DeValois & DeValois, 1991; see also Carrasco & Frieder, 1997), the eccentricity of visual targets might affect participants' target response latencies (e.g., Carrasco, Evert, Chang, & Katz, 1995; Wolfe, O'Neill, & Bennett, 1998; see also the later discussion). In order to ensure the spatial specificity of crossmodal cuing effects are not limited to a certain eccentricity, the two experiments reported here used different eccentricities of auditory cues and visual targets. The loudspeakers were evenly separated laterally by 30° in Experiment 1, and by 10° in Experiment 2. The loudspeakers and LEDs were vertically separated by approximately 19° in Experiment 1, and by 4° in Experiment 2 (see Figure 3).

---

<sup>1</sup> Kennett and Driver's (2014) study investigated how the within-hemifield alignment between hand postures (cue locations) and visual targets influenced between-hemifield exogenous spatial cuing effects. The cuing effects for the outer visual targets were larger when the participant's hands were aligned with those targets than when they were misaligned. However, cue-target alignment did not influence the spatial cuing effects for the inner visual targets. That is, RTs were faster when the cue and target stimuli were presented ipsilaterally than when they were presented contralaterally, regardless of the cue-target alignment. That said, their study did not investigate whether RTs on cued trials were any faster than when the cue and targets stimuli were presented within the same hemifield but from different locations.





If crossmodal exogenous spatial cuing effects are indeed spatially-specific, then (1) the presentation of a spatially-noninformative auditory cue would elicit faster RTs to visual targets presented from the same position as compared to those from a different lateral position within the cued hemifield (see Driver & Spence, 1998). In other words, within-hemifield spatial cuing effects should be observed. Furthermore, (2) RTs ought to be faster when the cue and target are presented from the same rather than different sides regardless of

the cue and target eccentricity. In other words, between-hemifield spatial cuing effects should be observed (e.g., Spence & Driver, 1997).

## EXPERIMENT 1

### Methods

#### *Participants*

Twenty-four participants (6 male and 18 female) volunteered to take part in this experiment through the Oxford Psychology Research participant recruitment scheme. Their average age was 27 years, ranging from 18 to 51 years. All of the participants reported normal or normal-to-corrected vision, and normal hearing. They were all right-handed except one left-hander based on their self-report. Experiment 1 took approximately 30 minutes to complete. The participants were paid £5 in return for taking part in the study. All of the studies reported here were approved by the Medical Sciences Interdivisional Research Ethics Committee at the University of Oxford (MS-IDREC-C1-2014-019), and were conducted according to the guidelines provided.

#### *Apparatus & materials*

The experiment was conducted in a darkened room (320 x 144 x 220cm), using MATLAB R2014a with PSYCHTOOLBOX 3.0.12 (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997). Four loudspeakers (Ricco 2.0 Channel Wooden Speaker Home Hifi System, model number: T2018) were situated in front of the participant's eye-level (see Figure 3). The centres of the loudspeaker cones at the outer-left and the outer-right positions were separated by 117cm, and the centre of each loudspeaker was separated from its nearest neighbour(s) by 39cm, evenly separated laterally by 30° seen from the participant's location between neighbouring loudspeakers. Each loudspeaker was equipped with a 3.5-inch diameter single-cone, capable of producing frequencies between 80Hz and 20kHz. Broadband noise was used as the auditory cue, matched at 72±1dBA between all four loudspeakers from the participant's head position. A pair of red target LEDs (12v 5mm with a luminance of 8000 millicandelas) was placed above and below each loudspeaker cone. The target LEDs were installed 20cm either above or below the centre of the loudspeaker cone, giving rise to approximately 38° of

vertical visual angle between the upper and lower targets from the participant's location (see Figure 3).

The four loudspeakers were placed on a wooden shelf, 117cm above the ground (as measured from the centre of the loudspeaker cones) to approximately match the participant's eye-level. They were connected to a Focusrite Scarlett 18i20 USB Computer Audio Interface (model number: AMS-SCARLETT-18i20). One red LED (identical to the target LEDs) was placed 58.5cm in front of the participant at eye-level, 117cm above the ground, between the two inner loudspeakers, as a fixation point. A computer monitor was placed below the frontal loudspeaker shelf in order to provide instructions. The participant was seated approximately 58.5cm from the red fixation LED.

### *Design*

There were two within-participants factors: Cue Location (outer-left, inner-left, inner-right, or outer-right) and Target Location (outer-left, inner-left, inner-right, or outer-right). Crossing these factors yielded 16 possible conditions, with each condition presented randomly six times within a block of trials. There were three blocks of trials, with the participants having the opportunity to take a short break after each block of trials.

### *Procedure*

Each trial started with the illumination of the fixation LED, followed by a random delay of between 400 and 650ms chosen from the discrete uniform distribution. Then, a task-irrelevant auditory cue was presented from one of the four loudspeakers for 100ms. The participants were instructed to ignore the auditory cues, which were followed, after a random delay of 100-200ms (stimulus onset asynchrony), by the presentation of a visual target from one of the eight possible locations for 140ms (see Figure 3). In order to indicate the elevation of the target LED, the participants had to press either the *up* or *down* arrow key on a computer keyboard placed on their lap, regardless of the lateral position from which the target was presented. The participants were encouraged to respond as rapidly and accurately as possible. When the participant's response had been registered, or else when 2,000ms had elapsed after the onset of the visual target, the fixation LED was extinguished and the next

trial was initiated. The instructions were displayed on the computer screen placed under the wooden shelf. Following the instructions, the participants completed ten practice trials. After the practice trials, the experimenter answered any questions and stepped out of the room before the actual trial began.

## Results

The data from two participant were excluded due to their failing to respond correctly to more than 95% of the trials. The average RT of each participant based on the total trials was compared to Tukey's (1977) outlier criteria<sup>2</sup>. A box-plot of the participants' average RTs revealed a median of 376ms, with 342ms as 25-percent quartile ( $Q_1$ ) and 444ms as 75-percent quartile ( $Q_3$ ). All of the participants' RTs were within the lower ( $Q_1 - 1.5 \text{ IQR}$ ) and upper limit ( $Q_3 + 1.5 \text{ IQR}$ ); no outliers were identified based on Tukey's method. Any trial data including incorrect responses, responses immediately following an incorrect response, and RTs outside of the range between 150 and 1,500ms were removed from the data analysis (see Spence & Driver, 1997, for similar exclusion criteria). A total of 201 trials (3.2%) was removed based on these exclusion criteria.

### *Reaction time analyses*

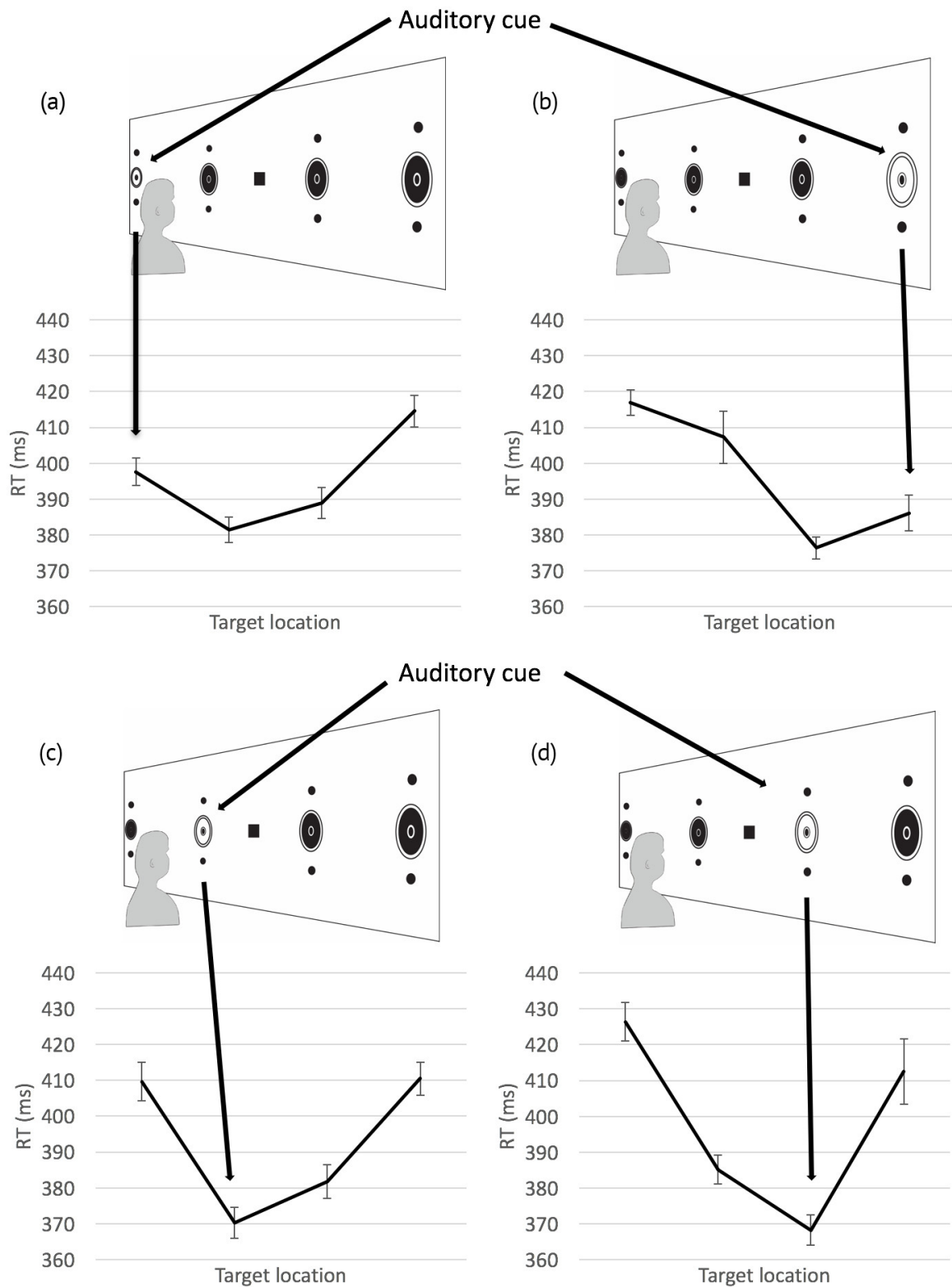
For comparison with Figure 2, the mean RTs in response to visual targets following the presentation of an auditory cue either from the outer-left, outer-right, inner-left, or inner-right position are summarised in Figure 4. The two within-participants factors of Cue Location and Target Location were entered into a RM-ANOVA. The analysis revealed a significant main effect of Target Location,  $F(3, 63) = 23.304$ ,  $MSE = 958.679$ ,  $p < .001$ ,  $\eta_p^2 = .526$ . The participants responded significantly more rapidly to the inner-left targets ( $M = 386\text{ms}$ ) than to the outer-left targets ( $M = 412\text{ms}$ ),  $t(21) = 5.653$ ,  $p < .001$ . Similarly, RTs to the inner-right targets ( $M = 379\text{ms}$ ) were significantly faster than those to the outer-right targets ( $M = 406\text{ms}$ ),  $t(21) = -5.463$ ,  $p < .001$ . The significant interaction between Cue and Target

---

<sup>2</sup> Based on Tukey's method, the lower limit was defined by the 25-percent quartile ( $Q_1$ ) subtracted by the interquartile range (IQR) multiplied by 1.5. The upper limit was defined by adding the 1.5 IQR to the 75-percent quartile ( $Q_3$ ). Data points falling outside the upper and lower limit were categorised as outliers.

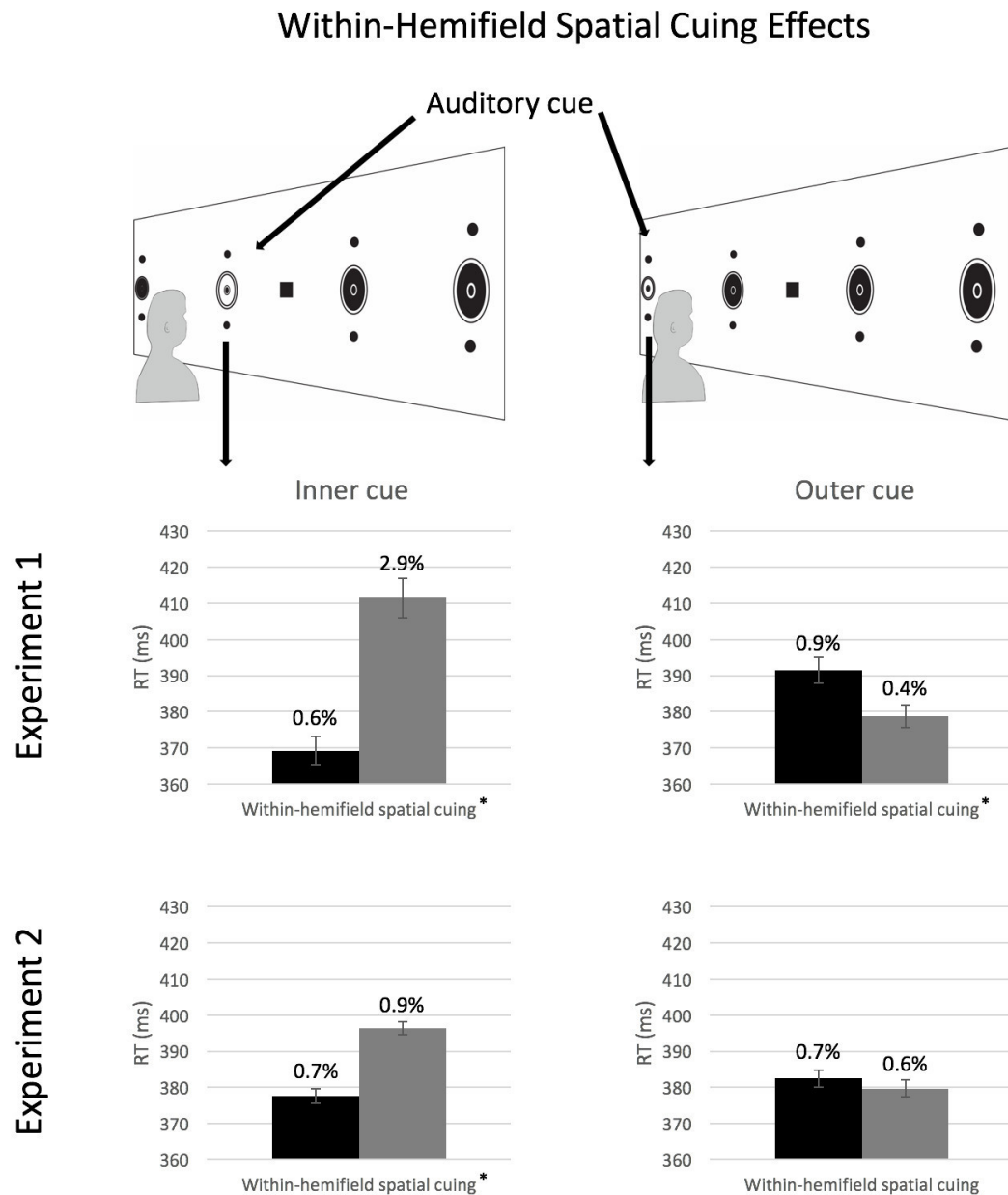
Location,  $F(9, 189) = 9.310$ ,  $MSE = 486.643$ ,  $p < .001$ ,  $\eta_p^2 = .307$ , highlights the presence of a significant crossmodal spatial cuing effect.

### Experiment 1



*Figure 4.* Mean reaction times (RTs) in milliseconds (ms) in the (a) outer-left, (b) outer-right, (c) inner-left, and (d) inner-right auditory cue conditions as a function of the Target Location in Experiment 1. The black dots and square represent the possible visual target locations and the fixation LED, respectively. The ellipses represent the loudspeakers, and the bright ellipse represents the location of the auditory cue. Standard errors of RTs are indicated by the vertical lines.

Two planned paired sample *t*-tests were conducted at each cue eccentricity with the RT data from those trials in which the cue and target were presented from the same hemifield. When the auditory cues were presented from the inner eccentricity, RTs were significantly faster when the visual targets were presented from the same position ( $M = 369\text{ms}$ ) than from a different lateral position within the cued hemifield ( $M = 411\text{ms}$ ),  $t(21) = -5.553$ ,  $p < .001$ . When the cues were presented from the outer eccentricity, however, RTs were significantly *slower* when the targets were presented from the same position ( $M = 391\text{ms}$ ) than from a different lateral position within the cued hemifield ( $M = 379\text{ms}$ ),  $t(21) = 2.620$ ,  $p = .016$ . In summary, the expected within-hemifield spatial cuing effect was documented only when the cues were presented from the inner eccentricity (see Figure 5).

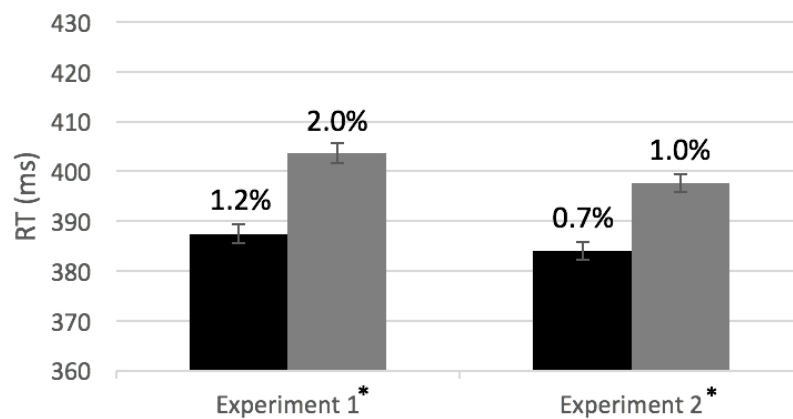


*Figure 5.* Mean reaction times (RTs) in milliseconds (ms) and error rates in percentages as a function of within-hemifield spatial cuing for Experiments 1 and 2. Black bars represent the mean RT on trials in which the cue and target were presented from the same position. Grey bars represent the mean RT on trials in which the cue and target were presented from different lateral locations within the cued hemifield. The left column represents the mean RT on trials in which the auditory cues were presented from the inner eccentricity regardless of the hemifield (left or right). The right column represents the mean RT on trials in which the cues were presented from the outer eccentricity regardless of the hemifield.

Standard errors of the mean RTs are indicated by the vertical lines. Asterisks are used to indicate significant RT difference between the within-hemifield spatial cuing conditions ( $\alpha = .05$ ).

A RM-ANOVA with the within-participants factor of Between-hemifield Spatial Cuing (cued if the cue and target were presented from the same side vs. uncued if not) was conducted in order to investigate whether there was a between-hemifield spatial cuing effect. The analysis revealed a significant main effect of Between-hemifield Spatial Cuing,  $F(1, 21) = 35.614$ ,  $MSE = 81.105$ ,  $p < .001$ ,  $\eta_p^2 = .629$ , with participants responding more rapidly to targets when the cue was presented from the same ( $M = 388\text{ms}$ ) rather than the opposite ( $M = 404\text{ms}$ ) side (see Figure 6).

### Between-Hemifield Spatial Cuing Effects



*Figure 6.* Mean reaction times (RTs) in milliseconds (ms) and error rates in percentages as a function of between-hemifield spatial cuing for Experiments 1 and 2. Black bars represent the mean RT on those trials in which the cue and target were presented in the same hemifield. Grey bars represent the mean RT on those trials in which the cue and target were presented from different hemifields. Standard errors of mean RTs are indicated by the vertical lines. Asterisks are used to indicate significant between-hemifield spatial cuing effects in the RT data ( $\alpha = .05$ ).



*Error rate analyses*

A RM-ANOVA with the two within-participants factors of Cue Location and Target Location revealed a significant main effect of Target Location,  $F(3, 63) = 7.464$ ,  $MSE = .001$ ,  $p < .001$ ,  $\eta_p^2 = .262$ . The participants made significantly more errors with the outer-left visual targets ( $M = 2.3\%$ ) than with the inner-left targets ( $M = 0.9\%$ ),  $t(21) = 3.334$ ,  $p = .003$ . Furthermore, the participants made significantly more errors when targets were presented from the outer-right position ( $M = 2.5\%$ ) than when presented from the inner-right position ( $M = 0.6\%$ ),  $t(21) = -2.871$ ,  $p = .009$ . The average ER was 1.6%, with the maximum ER of 3.8% and the minimum ER of 0.0%.

**Discussion**

The results of Experiment 1 confirmed the existence of the standard between-hemifield spatial cuing effect (e.g., Spence & Driver, 1997). More importantly, the hypothesised within-hemifield spatial cuing effect was only documented when the auditory cues were presented from the inner eccentricity, thus contrasting with the conventional view concerning the spatial specificity of exogenous crossmodal cuing effects (e.g., Driver & Spence, 1998; Spence et al., 2004). What is more, following the presentation of an auditory cue from the outer eccentricity, RTs to targets were significantly *slower* when the cue and target were presented from the same position than when the stimuli were presented from different lateral positions within the hemifield. Based on the lack of any interaction between Cue Location and Target Location in the error data, the documented spatial cuing effects would appear to reflect a genuine attentional facilitation rather than any kind of speed-accuracy trade-off (see Duncan, 1980; Spence, Pavani, & Driver, 2000).

The reversed within-hemifield spatial cuing effect following the presentation of the auditory cues from the outer eccentricity presumably reflects the confounding influence of *eccentricity* on participants' visual target discrimination performance (e.g., Carrasco et al., 1995; Wolfe et al., 1998). Remember here that those visual stimuli that are presented from, or near to, fixation tend to be detected more rapidly (and accurately) than those that are presented from further away. In Experiment 1, the inner and outer eccentricities were separated laterally by 30°. Consequently, the inner targets should have been responded to more rapidly than the outer targets, as confirmed here, regardless of the location of the auditory cue (see Carrasco

et al., 1995; Wolfe et al., 1998). Therefore, any crossmodal facilitation effects for the outer targets from the presentation of an auditory cue at the same position might have not been detected due to the slower baseline RTs to outer targets than those to inner targets. Similarly, the confirmed within-hemifield spatial cuing effect following the presentation of the auditory cues from the inner eccentricity can also be explained by the eccentricity effect alone.

Generally-speaking, the more eccentric the visual stimuli, the more pronounced the eccentricity effect becomes (Carrasco et al., 1995; see also Carrasco & Frieder, 1997). In an attempt to reduce the influence of this effect in the next experiment, all of the stimuli were now presented within 20° laterally from the fixation to potentially reduce the lateral eccentricity confound on spatial cuing (Carrasco et al., 1995; Carrasco & Frieder, 1997). One might suspect that the magnitude of spatial cuing effects would decrease due to the spatial proximity of the stimuli (e.g., the closer lateral distances between stimuli than those in Experiment 1). However, the evidence suggests that humans can discriminate the relative locations of two auditory cues when they are laterally separated by 2° or less (see Perrott, 1993). Given that the inner and outer locations will be separated by a visual angle of 10°, we expected that the results of Experiment 2 would be similar to those already reported in Experiment 1.

## EXPERIMENT 2

### Methods

#### *Participants*

Twenty participants from the Oxford Psychology Research participant recruitment scheme took part in this experiment (4 male and 16 female) with their average age of 26 years, ranging from 19 to 35 years. All of the participants self-reported normal or normal-to-corrected vision, and normal hearing. They were all right-handed. Experiment 2 took approximately 30 minutes. All of the participants were paid £5 at the end of the experiment.

#### *Apparatus & materials*

These were identical to those in Experiment 1 with the following exceptions: (1) The participants' eyes were approximately 77cm away from the fixation LED; (2) all four loudspeakers were placed next to each other, creating a 40° auditory cue range between the left wall of the far left loudspeaker and the right wall of the far right loudspeaker (56cm); (3) two target LEDs were installed 5.5cm above and below the centre of each loudspeaker cone, with approximately 8° of the vertical viewing angle. The inner and outer visual targets were placed approximately 5° (7cm) and 15° (21cm) from the fixation LED on each side (see Figure 3).

### *Design, procedure, & data analyses*

These were the same as in Experiment 1.

## **Results**

All of the participants responded to 100% of the trials. One participant failed to respond correctly to more than 95% of the trials, and therefore his/her data was excluded from the analyses. A box-plot of the participants' average RTs of the total trials revealed a median of 386ms, with 351ms and 427ms as  $Q_1$  and  $Q_3$ , respectively. The RTs of all participants were within the lower and upper limit based on Tukey's (1977) method; no outliers were identified. A total of 98 trials (1.8%) were removed from the data analyses based on the error response exclusion criteria used in Experiment 1.

### *Reaction time analyses*

The mean RTs to visual targets following the presentation of an auditory cue either from the outer-left, outer-right, inner-left, or inner-right position are summarised in Figure 7. The within-participants factors of Cue Location and Target Location were entered into a RM-ANOVA. The analysis revealed a significant main effect of Target Location,  $F(3, 54) = 18.984$ ,  $MSE = 382.509$ ,  $p < .001$ ,  $\eta_p^2 = .513$ . Once again, the participants responded significantly more rapidly to the targets presented from the inner-left position ( $M = 385\text{ms}$ ) than from the outer-left position ( $M = 396\text{ms}$ ),  $t(18) = 4.141$ ,  $p = .001$ . RTs to the targets

presented from the inner-right ( $M = 381\text{ms}$ ) were also significantly faster than those from the outer-right ( $M = 402\text{ms}$ ),  $t(18) = -8.117$ ,  $p < .001$ . There was also a significant interaction between Cue and Target Location,  $F(9, 162) = 9.515$ ,  $MSE = 255.721$ ,  $p < .001$ ,  $\eta_p^2 = .346$ , indicating the presence of a significant audiovisual spatial cuing effect.

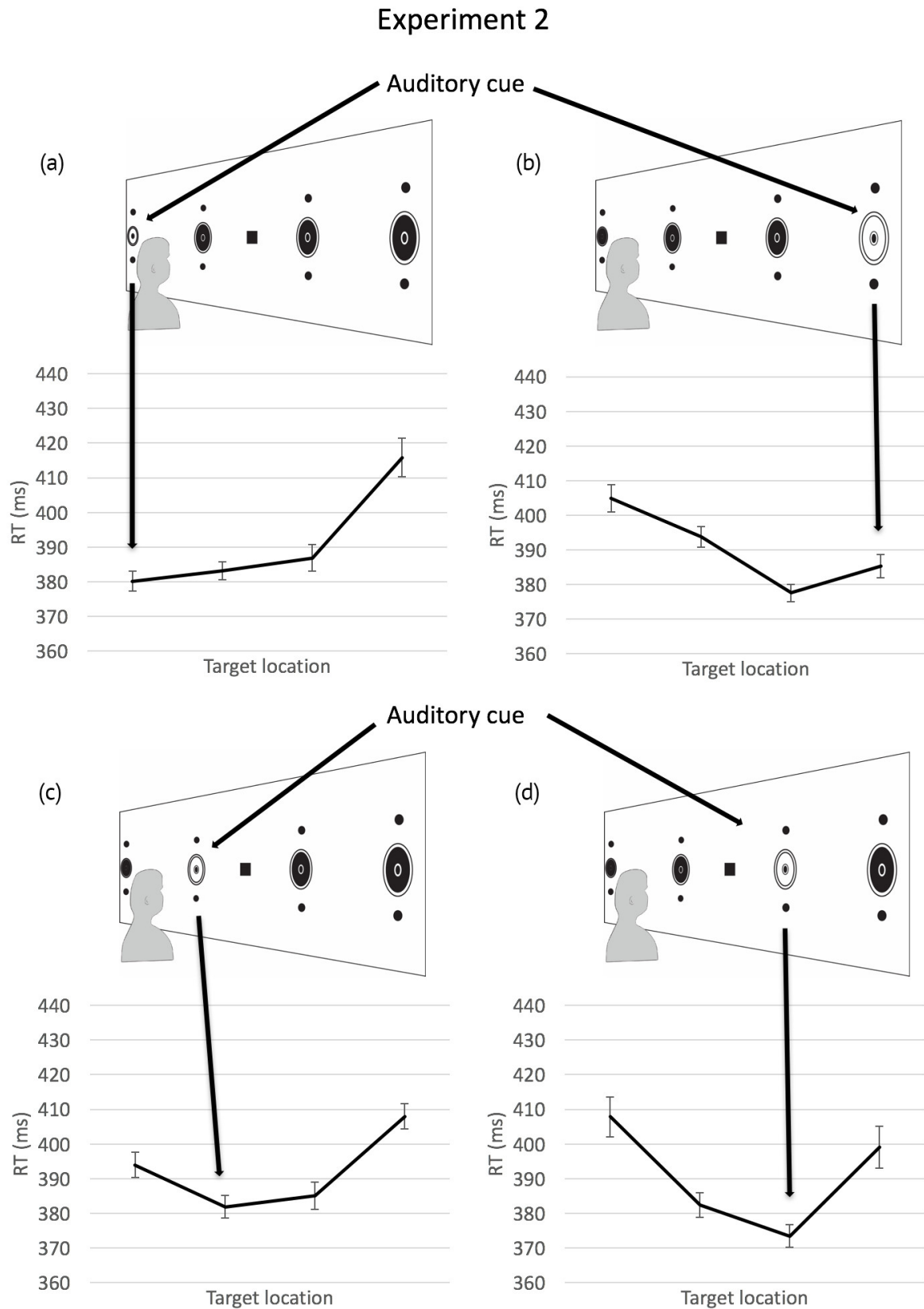


Figure 7. Mean reaction times (RTs) in milliseconds (ms) for Cue Location at the (a) outer-left, (b) outer-right, (c) inner-left, and (d) inner-

right as a function of the Target Location in Experiment 2. The black dots and square represent the visual LEDs and the fixation LED, respectively. The ellipses represent the loudspeaker locations, and the bright one represents the auditory cue location. Standard errors of RTs are indicated using the vertical lines.

Two planned paired sample *t*-tests were conducted in order to test for within-hemifield spatial cuing effects. The analyses revealed that when the auditory cues were presented from the inner eccentricity, RTs were significantly faster when the visual targets were presented from the same position ( $M = 378\text{ms}$ ) than from a different lateral position within the cued hemifield ( $M = 396\text{ms}$ ),  $t(18) = -7.840$ ,  $p < .001$ . When the cues were presented from the outer eccentricity, however, RTs to targets presented from the same position ( $M = 383\text{ms}$ ) were not statistically different compared to those presented from a different lateral position within the cued hemifield ( $M = 380\text{ms}$ ),  $t(18) = .829$ ,  $p = .418$ . Once again, the expected within-hemifield spatial cuing effect was documented only when the cues were presented from the inner eccentricity (see Figure 5). A RM-ANOVA with the within-participants factor of Between-hemifield Spatial Cuing revealed a significant main effect of Between-hemifield Spatial Cuing,  $F(1, 18) = 28.858$ ,  $MSE = 60.585$ ,  $p < .001$ ,  $\eta_p^2 = .616$ , with participants responding significantly more rapidly to targets when the cue was presented from the same ( $M = 384\text{ms}$ ) rather than the opposite ( $M = 398\text{ms}$ ) hemifield (see Figure 6).

#### *Error rate analyses*

The RM-ANOVA with the two within-participants factors of Cue Location and Target Location did not reveal any significant terms. The mean ER in any single condition was 0.9%, with the maximum ER of 2.0% and the minimum ER of 0.3%.

#### **Discussion**

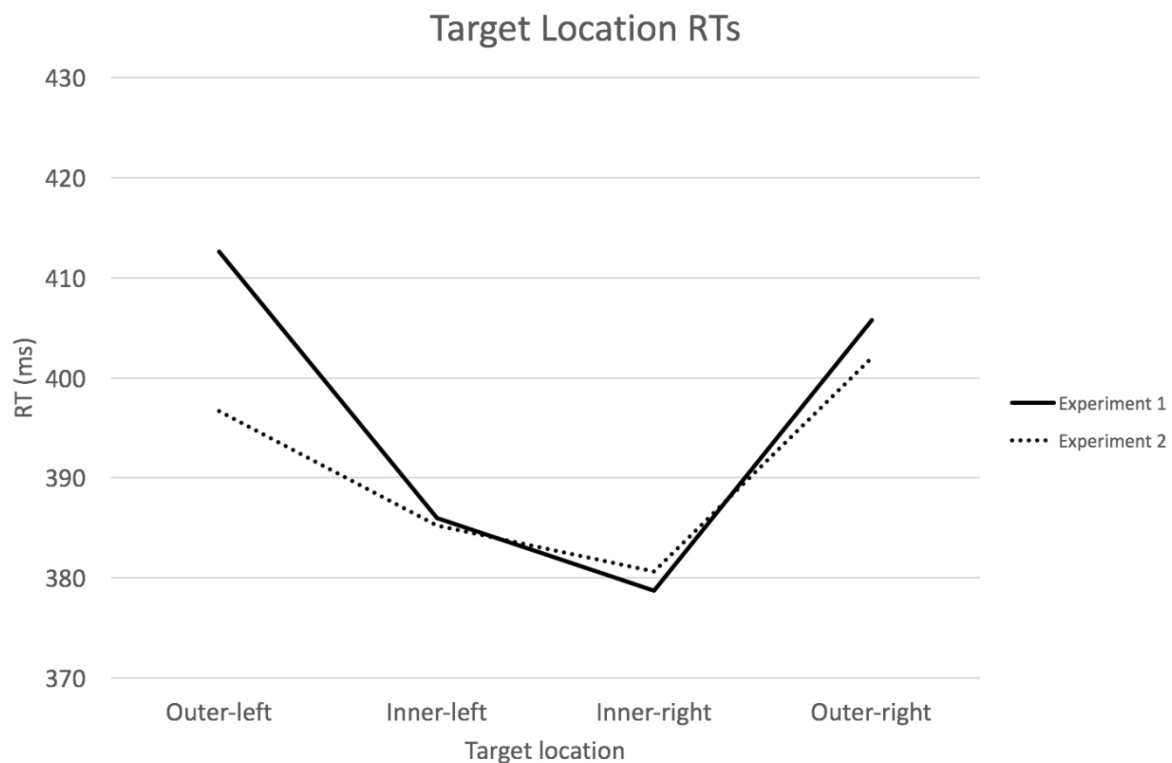
As in Experiment 1, the results of Experiment 2 confirmed the standard between-hemifield spatial cuing effect. The findings of the within-hemifield spatial cuing effect are similar to

those reported in Experiment 1. Following the presentation of an auditory cue from the inner eccentricity, the participants responded significantly more rapidly to the targets presented from the same position than to those from a different lateral position within the cued hemifield. However, no such within-hemifield spatial cuing effect was found when the auditory cues were presented from the outer eccentricity.

It is noteworthy that RTs to the inner targets were significantly faster than those to the outer targets no matter where the auditory cues were presented from. In fact, the participants in both experiments responded significantly more rapidly to the inner targets than to the outer targets, regardless of the spatial cuing conditions. Therefore, even though the small experiment set-size should have, at least in theory, mitigated the eccentricity effect, the data in Experiment 2 suggests that the documented within-hemifield cuing effects cannot be explained by the spatially-specific attentional shift following the presentation of an auditory cue.

### **General Discussion**

The results of the present study add further weight to the suggestion that the presentation of a spatially-noninformative, task-irrelevant auditory cue can facilitate the perception of visual targets presented from the same rather than opposite hemifield. However, the results of the two experiments reported here provide evidence against the conventional view concerning the spatial specificity of crossmodal audiovisual cuing effects, particularly from those trials in which the auditory cues were presented from the outer eccentricity. Furthermore, both experiments demonstrated that RTs to the inner targets were significantly faster than those to the outer targets regardless of the cue location, thus confirming the existence of the eccentricity confound (see Figure 8). Therefore, the documented within-hemifield spatial cuing effects on those trials in which the auditory cues were presented from the inner locations could simply reflect an eccentricity effect.



*Figure 8.* Mean reaction times (RTs) in milliseconds (ms) in Experiments (solid-line) 1 and (dotted-line) 2 as a function of Target Location.

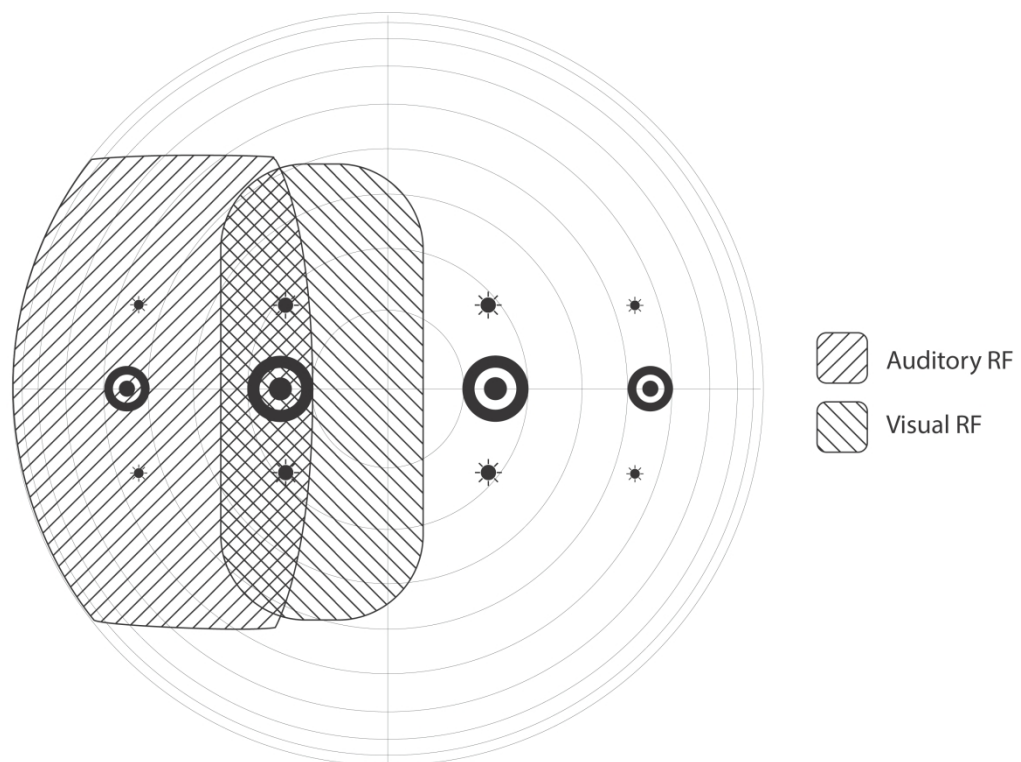
One potential factor of the eccentricity confound could be the difference in spatial resolution between the fovea and the periphery in retina (see DeValois & DeValois, 1991). One of the major areas that ganglion cells project from retina is to lateral geniculate nucleus (and also to the superior colliculus [SC]), which then relays the information to primary visual cortex<sup>3</sup> (V1; DeValois & DeValois, 1991). However, ganglion cell density decreases as a function of the retinal eccentricity (from the fovea; Fischer, 1973). When the retinal projection finally arrives at V1, the cortical map reflects the logarithmic magnification of the retinal representation (see Tootell, Silverman, Switkes, & De Valois, 1982). Indeed, Carrasco and Frieder (1997) demonstrated in their visual search study that the eccentricity effect could be

<sup>3</sup> It is worth noting that the SC receives inputs from ganglion cells as well as from various cortical areas including primary visual cortex (V1; see Collins, Lyon, & Kaas, 2005). As a result, the cortical magnification in V1 could potentially influence how the multisensory neurons in the SC respond to the visual targets in audiovisual cuing studies.



eliminated if the stimuli on the screen were scaled by the cortical magnification factor (CMF; Rovamo & Virsu, 1979; Virsu & Rovamo, 1979).

Evidence from animal research suggests that the properties of the SC might also be attributable to the documented within-hemifield spatial cuing effects in the current study. For instance, based on data from anaesthetised cats, Meredith and Stein (1996) reported that the RFs of audiovisual SC neurons responsive to the frontal space exhibited visual RFs aligned centrally within the auditory RFs. However, for neurons with RFs falling in lateral space, the visual RFs tended to be pushed close to, or over the medial border of auditory RFs (Meredith & Stein, 1996; see also Wallace & Stein, 2007). Therefore, it is possible to elicit response enhancement even when the auditory cues were presented from the outer location and visual targets from the inner location in the same hemifield, but not the other way around (see Figure 9).



*Figure 9.* Schematic RF spatial registry of a SC single audiovisual neuron found in the peripheral sensory space. The horizontal and vertical lines represent 0° azimuth and 0° elevation, respectively, and each concentric circle 10° of space. Note that the auditory RF, although maintaining a considerable overlap with the visual RF, is located toward the peripheral

sensory space. The illustration is based on the figures reported in Meredith and Stein (1996; also see Kadunce et al., 2001).

Neurophysiological studies in animals further suggest that multisensory neurons in the SC can vary widely in size depending on their locations. For example, the multisensory neurons in the rostral portion of the cat SC, responsive to stimuli presented from frontal space, have been reported to have RFs often less than 10° for visual and 20° for auditory (in diameter; see Kadunce, Vaughan, Wallace, Benedek, & Stein, 2001). On the other hand, multisensory neurons in the caudal SC that respond to stimuli presented from the periphery have visual RFs ranging from 40 to 100° and auditory RFs from 60 to 135° in diameter (see Kadunce et al., 2001). Therefore, the attentional shift from the centre elicited by the auditory cues from the inner locations could be potentially more spatially-specific than from the outer locations, as demonstrated in both experiments.<sup>4</sup>

---

<sup>4</sup> Since the data from the current study was collected from human participants, it is somewhat speculative to relate our results to the neurophysiological findings of the SC in animals. However, the spatial properties of multisensory neurons are not limited only to cats; the SC neurons of monkeys also share similar properties such as topographic alignments of RFs with emphasis on visual space and the spatial register of RFs governing multisensory integration (Wallace, Wilkinson, & Stein, 1996). Furthermore, the neurophysiological data provides an additional theoretical explanation for the results reported in the current study. In conclusion, at least in theory, the findings on within-hemifield spatial cuing effects could be attributable to the correspondence patterns of audiovisual RFs in the SC, the eccentricity effect, or both.

The results from the two experiments reported here suggest that the presentation of an auditory cue does not always elicit faster RTs to visual targets presented from the same position rather than from a different lateral position within the cued hemifield. Regarding the spatial specificity of within-hemifield cuing effects, there are two questions remaining that are undoubtedly worthy of investigation. The first question is: Are *cuing benefits*, calculated by subtracting the current RTs from the baseline RT to targets at each eccentricity, spatially-specific? If audiovisual cuing effects without the eccentricity confound are, in fact, spatially-specific, the cuing benefits should be larger when the cue and target stimuli are presented from the same, rather than from different lateral positions within the cued hemifield. The second question is: Can the target perception be facilitated maximally following the presentation of an auditory cue from the same position rather than from a different position within the target hemifield? It is worth pausing for a moment here to note that the spatial specificity of within-hemifield spatial cuing effects has been understood based on how the presentation of an auditory cue from a given position would modulate RTs to visual targets from the same position as compared to those from a different lateral position *within the cued hemifield* (see Driver & Spence, 1998; Spence et al., 2004; Spence & McDonald, 2004). On the other hand, the latter question focuses on how the RTs to visual targets would be modulated by the presentation of an auditory cue either from the same position or from a different lateral position *within the target hemifield*. Both investigations would enable us to study the spatial specificity of within-hemifield cuing effects without the target eccentricity confound.

In order to test the spatially-specific cuing benefit hypothesis, cuing benefits were calculated for each participant based on his/her average RT to targets at each eccentricity. Planned paired sample *t*-tests for the RT data from Experiment 1 revealed that when the auditory cues were presented from the inner eccentricity, cuing benefits were significantly larger when the visual targets were presented from the same position ( $M = 13\text{ms}$ ) than from a different lateral position within the cued hemifield ( $M = -3\text{ms}$ ),  $t(21) = 3.143$ ,  $p = .005$ . When the cues were presented from the outer eccentricity, cuing benefits were also significantly larger when the targets were presented from the same position ( $M = 18\text{ms}$ ) than from a different lateral position within the cued hemifield ( $M = 4\text{ms}$ ),  $t(21) = 2.712$ ,  $p = .013$ . Another set of planned paired sample *t*-tests for Experiment 2 revealed that when the cues were presented from the inner eccentricity, there was a non-significant trend for the cuing benefits to be larger when the targets were presented from the same position ( $M = 5\text{ms}$ ) than from a different lateral

position within the cued hemifield ( $M = 3\text{ms}$ ),  $t(18) = 1.360$ ,  $p = .190$ . When the cues were presented from the outer eccentricity, cuing benefits were significantly larger when the targets were presented from the same position ( $M = 17\text{ms}$ ) than from a different lateral position within the cued hemifield ( $M = 3\text{ms}$ ),  $t(18) = 3.979$ ,  $p = .001$ .

Two planned paired sample  $t$ -tests were conducted to investigate whether the perception of visual targets can be facilitated maximally when the auditory cues were presented from the same position rather than from a different lateral position within the target hemifield. Planned paired sample  $t$ -tests for the RT data from Experiment 1 revealed that RTs to inner targets were faster when the cues were presented from the same position ( $M = 369\text{ms}$ ) than from a different lateral position within the target hemifield ( $M = 379\text{ms}$ ),  $t(21) = -2.260$ ,  $p = .035$ . Similarly, RTs to outer targets were faster when the cues were presented from the same position ( $M = 391\text{ms}$ ) than from a different lateral position within the target hemifield ( $M = 411\text{ms}$ ),  $t(21) = -3.065$ ,  $p = .006$ . The same analyses for Experiment 2 revealed that RTs to inner targets were faster (although not significantly) when the cues were presented from the same position ( $M = 378\text{ms}$ ) than from a different lateral position within the target hemifield ( $M = 380\text{ms}$ ),  $t(18) = -.647$ ,  $p = .526$ . RTs to the outer targets were significantly faster when the cues were presented from the same position ( $M = 383\text{ms}$ ) than from a different lateral position within the target hemifield ( $M = 396\text{ms}$ ),  $t(18) = -4.668$ ,  $p < .001$ .

The re-analyses revealed that the magnitude of response facilitation for visual targets was statistically maximum when auditory cues are presented from the same position, except when the lateral eccentricity of the cue and target is  $5^\circ$  from the fixation on each side. The lack of significant within-hemifield spatial cuing benefit/effect in Experiment 2 from the inner eccentricity suggests that there is little advantage of auditory spatial cuing for visual targets close to the current direction of gaze. Furthermore, the magnitude of cuing benefit/effect was larger from the outer eccentricity than from the inner eccentricity in both experiments. In summary, the presentation of a spatially-noninformative auditory cue *does not* always lead to faster RTs to visual targets presented from the same position as compared to those from a different lateral position within the cued hemifield. However, the perception of visual targets is maximally facilitated when auditory cues are presented from the same position regardless of the cue-target eccentricity.

Taken together, the results of both experiments reported here add further support for the existence of spatial cuing effects between hemifields. Additionally, the present study

provides the first empirical evidence on the spatial specificity of exogenous audiovisual cuing effects within the cued hemifield without a potential response bias confound. These findings may have real-world applications, such as developing auditory crash avoidance systems (CAS; e.g., Lee, Olsen, & Wierwille, 2004). Between-hemifield spatial auditory cuing should be able to facilitate response to visual targets (e.g., potential dangers on the road) on the cued side. However, if even greater facilitation of the response to a target is desired, an auditory warning cue should be presented from a specific location that is in the direction of a visual target (e.g., Ho & Spence, 2005). If the CAS can accommodate only one auditory alarm position on each side, vehicle operators would benefit more if the warnings were designed to assist them to be aware of visual targets presented further away from, rather than close to, the central gaze.

## REFERENCES

- Baldwin, C. L., Spence, C., Bliss, J. P., Brill, J. C., Wogalter, M. S., Mayhorn, C. B., & Ferris, T. K. (2012). Multimodal cueing: The relative benefits of the auditory, visual, and tactile channels in complex environments. *Proceedings of the 56<sup>th</sup> Human Factors and Ergonomics Society meeting*, 56, 1431-1435.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10, 433-436.
- Carrasco, M., Evert, D. L., Chang, I., & Katz, S. M. (1995). The eccentricity effect: Target eccentricity affects performance on conjunction searches. *Perception & Psychophysics*, 57(8), 1241-1261.
- Carrasco, M., & Frieder, K. S. (1997). Cortical magnification neutralizes the eccentricity effect in visual search. *Vision Research*, 37(1), 63-82.
- Collins, C. E., Lyon, D. C., & Kaas, J. H. (2005). Distribution across cortical areas of neurons projecting to the superior colliculus in new world monkeys. *The Anatomical Record Part A: Discoveries in Molecular, Cellular, and Evolutionary Biology*, 285(1), 619-627.
- DeValois, R. L., & DeValois, K. K. (1991). *Spatial vision*. New York, NY: Oxford University Press.
- Downing, C. J. (1988). Expectancy and visual-spatial attention: Effects on perceptual quality. *Journal of Experimental Psychology: Human Perception and Performance*, 14(2), 188-202.
- Downing, C. J., & Pinker, S. (1985). The spatial structure of visual attention. In M. I. Posner & O. S. M. Marin (Eds.), *Attention and Performance XI* (pp. 171-187). Hillsdale, NJ: Erlbaum.
- Driver, J., & Spence, C. (1998). Cross-modal links in spatial attention. *Philosophical Transactions of the Royal Society Series B*, 353(1373), 1319-1331.
- Duncan, J. (1980). The demonstration of capacity limitation. *Cognitive Psychology*, 12(1),

75-96.

- Fischer, B. (1973). Overlap of receptive field centers and representation of the visual field in the cat's optic tract. *Vision Research*, 13(11), 2113-2120.
- Ho, C., & Spence, C. (2005). Assessing the effectiveness of various auditory cues in capturing a driver's visual attention. *Journal of Experimental Psychology: Applied*, 11(3), 157-174.
- Ho, C., & Spence, C. (2008). *The multisensory driver: Implications for ergonomic car interface design*. Aldershot, HA: Ashgate.
- Ho, C., Tan, H. Z., & Spence, C. (2006). The differential effect of vibrotactile and auditory cues on visual spatial attention. *Ergonomics*, 49(7), 724-738.
- Kadunce, D. C., Vaughan, W. J., Wallace, M. T., & Stein, B. E. (2001). The influence of visual and auditory receptive field organization on multisensory integration in the superior colliculus. *Experimental Brain Research*, 139(3), 303-310.
- Kennett, S., & Driver, J. (2014). Within-hemifield posture changes affect tactile-visual exogenous spatial cueing without spatial precision, especially in the dark. *Attention, Perception, & Psychophysics*, 76(4), 1121-1135.
- Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception*, 36(14), 1-89.
- Lee, S., Olsen, E., & Wierwille, W. (2004). A comprehensive examination of naturalistic lane-changes. *U.S. Department of Transportation, DOT HS 809 702*.
- Mangun, G. R., & Hillyard, S. A. (1988). Spatial gradients of visual attention: Behavioral and electrophysiological evidence. *Electroencephalography and Clinical Neurophysiology*, 70(5), 417-428.
- Meredith, M. A., & Stein, B. E. (1996). Spatial determinants of multisensory integration in cat superior colliculus neurons. *Journal of Neurophysiology*, 75(5), 1843-1857.
- Mock, J. R., Seay, M. J., Charney, D. R., Holmes, J. L., & Golob, E. J. (2015). Rapid cortical dynamics associated with auditory spatial attention gradients. *Frontiers in Neuroscience*, 9(179):1-14.

- Mondor, T. A., & Zatorre, R. J. (1995). Shifting and focusing auditory spatial attention. *Journal of Experimental Psychology: Human Perception and Performance*, 21(2), 387-409.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10(4), 437-442.
- Perrott, D. R. (1993). Auditory and visual localization: Two modalities and one world. *Proceedings of the Auditory Engineering Society 12<sup>th</sup> International Conference* (pp. 221-231). Snekkersten, Copenhagen
- Rovamo, J., & Virsu, V. (1979). An estimation and application of the human cortical magnification factor. *Experimental Brain Research*, 37(3), 495-510.
- Schmitt, M., Postma, A., & De Haan, E. H. (2001). Cross-modal exogenous attention and distance effects in vision and hearing. *European Journal of Cognitive Psychology*, 13(3), 343-368.
- Shulman, G. L., Wilson, J., & Sheehy, J. B. (1985). Spatial determinants of the distribution of attention. *Perception & Psychophysics*, 37(1), 59-65.
- Spence, C. (J.), & Driver, J. (1994). Covert spatial orienting in audition: Exogenous and endogenous mechanisms. *Journal of Experimental Psychology: Human Perception and Performance*, 20(3), 555-574.
- Spence, C. (J.), & Driver, J. (1996). Audiovisual links in endogenous covert spatial attention. *Journal of Experimental Psychology: Human Perception and Performance*, 22(4), 1005-1030.
- Spence, C., & Driver, J. (1997). Audiovisual links in exogenous covert spatial orienting. *Perception & Psychophysics*, 59(1), 1-22.
- Spence, C., & McDonald, J. (2004). The cross-modal consequences of the exogenous spatial orienting of attention. In G. A. Calvert, C. Spence, & B. E. Stein (Eds.), *The handbook of multisensory processing* (pp. 3-25). Cambridge, MA: MIT Press.
- Spence, C., McDonald, J., & Driver, J. (2004). Exogenous spatial cuing studies of human crossmodal attention and multisensory integration. In C. Spence & J. Driver (Eds.), *Crossmodal space and crossmodal attention* (pp. 277-320). Oxford, UK: Oxford



University Press.

- Spence, C., Pavani, F., & Driver, J. (2000). Crossmodal links between vision and touch in covert endogenous spatial attention. *Journal of Experimental Psychology: Human Perception and Performance*, 26(4), 1298-1319.
- Teder-Sälejärvi, W. A., & Hillyard, S. A. (1998). The gradient of spatial auditory attention in free field: An event-related potential study. *Perception & Psychophysics*, 60(7), 1228-1242.
- Tootell, R. B. H., Silverman, M. S., Switkes, E., & De Valois, R. L. (1982). Deoxyglucose analysis of retinotopic organization in primate striate cortex. *Science*, 218(4575), 902-904.
- Tukey, J. W. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley.
- Virsu, V., & Rovamo, J. (1979). Visual resolution, contrast sensitivity, and the cortical magnification factor. *Experimental Brain Research*, 37(3), 475-494.
- Wallace, M. T., & Stein, B. E. (2007). Early experience determines how the senses will interact. *Journal of Neurophysiology*, 97(1), 921-926.
- Wallace, M. T., Wilkinson, L. K., & Stein, B. E. (1996). Representation and integration of multiple sensory inputs in primate superior colliculus. *Journal of Neurophysiology*, 76(2), 1246-1266.
- Wolfe, J. M., O'Neill, P., & Bennett, S. C. (1998). Why are there eccentricity effects in visual search? Visual and attentional hypotheses. *Perception & Psychophysics*, 60(1), 140-156.