Journal of Eye Movement Research
2(3):4,1-13

Cueing Visual Attention to Spatial Locations With Auditory Cues

Matthew Kean
University of Manchester

Trevor J. Crawford
Lancaster University

We investigated exogenous and endogenous orienting of visual attention to the spatial location of an auditory cue. In Experiment 1, significantly faster saccades were observed to visual targets appearing ipsilateral, compared to contralateral, to the peripherally-presented cue. This advantage was greatest in an 80% target-at-cue (TAC) condition but equivalent in 20% and 50% TAC conditions. In Experiment 2, participants maintained central fixation while making an elevation judgment of the peripheral visual target. Performance was significantly better for the cued side of the display, and this advantage was equivalent across the three expectancy conditions. Results point to attentional processes, rather than simply ipsilateral response preparation, and suggest that orienting visual attention to a sudden auditory stimulus is difficult to avoid.

Keywords: Auditory cues, cueing, covert, overt, saccade, spatial attention

Introduction

Posner (1980) introduced the spatial cueing paradigm for investigating the allocation of attention to peripheral locations. In this task, a cue precedes the onset of a target stimulus, to which a response is required. In this paradigm, a cued or 'valid' target location typically enjoys processing advantages relative to an uncued or 'invalid' location. The processing advantages may be, for example, faster detection of the target, or more accurate discrimination of the target. Such observable processing advantages are thought to be indicative of an attention shift to the cued location (e.g., Posner, 1980; Bonnel *et al.*, 1987). Essentially, when a peripheral cue is presented, attention is reflexively drawn to its location (Jonides, 1981). This phenomenon, known as attention capture (or attentional capture), acts very rapidly, exerting its maximal influence on attentional orientation 100 msec after cue onset (Cheal & Lyon, 1991; Wright & Ward, 1994). If the target appears nearby, and very soon afterwards, a response can occur relatively rapidly, or accurately in the case of target discrimination. If the target appears in another location, attention must be disengaged from the cue, then moved to – and engaged upon – the other location, before the response may be executed (Posner, 1980). Slower latency or poorer discrimination for invalid trials reflects this delay in moving – or 'reorienting' – attention away from the cued location. A related phenomenon, referred to as oculomotor capture, occurs when a peripheral event causes a reflexive saccade to its location (Theeuwes *et al.*, 1998).

A distinction is commonly drawn between two fundamentally different processes that are capable of effecting attention shifts; one requiring conscious direction (i.e., a top-down, voluntary process), and another that acts independently of conscious control (i.e., a reflexive, bottom-up, or stimulus-driven process). These two processes are known as the endogenous and exogenous orienting mechanisms, respectively (Posner, 1980; Müller & Rabbitt, 1989), and this distinction is supported by evidence of different neural underpinnings associated with the two different processes (e.g., Corbetta *et al.*, 1993; Robertson & Rafal, 2000; Robinson & Kertzman, 1995; Rosen *et al.*, 1999; Zackon *et al.*, 1999).

In the early days of attention research, the predominant view concerning the cognitive processes involved in the orienting of attention was that attention can be shifted

DOI 10.16910/jemr.2.3.4
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

either reflexively *or* voluntarily from one location to another. This dichotomous notion that attention is either exogenously captured or endogenously directed has been called into question. For instance, attention capture by a peripherally presented visual stimulus was originally thought to be unavoidable (Jonides, 1981), but some findings have indicated that attention capture is – at least under some circumstances – subject to top-down (i.e., consciously directed) influence (Yantis & Jonides, 1990; Folk *et al.*, 1992; Tepin & Dark, 1992). Contemporary models depict attentional control – at least in the realm of visual attention – as the product of a dynamic interaction between stimulus properties and the expectations and goals of the observer (Yantis, 1998), or, in other words, between the exogenous and endogenous orienting mechanisms.

We used an adapted version of Posner's (1980) cueing paradigm to measure shifts of visual attention to the spatial location of an auditory stimulus. Several researchers have used this adapted paradigm (e.g., Buchtel & Butter, 1988; Spence & Driver, 1997; Schmitt *et al.*, 2000), where an auditory cue precedes a visual target stimulus, either in a nearby or distant location. Using this method of cross-modal cueing, we sought to investigate possible cross-modal links in the orienting of exogenous and endogenous attention.

It is known that auditory attention is reflexively drawn to the spatial location of an auditory stimulus (Spence & Driver, 1994). However, interesting links occur between sensory modalities, such as the observation that saccadic trajectories tend to deviate slightly away from the spatial locations of auditory and tactile distractors (see: Walker & McSorley, this volume). With respect to cross-modal attention, there is evidence for the reflexive capture of visual attention by an auditory stimulus to the same spatial location (Spence & Driver, 1997; Schmitt *et al.*, 2000; Schmitt *et al.*, 2001; Mazza *et al.,* 2007), and it has also been shown that visual attention may be purely endogenously directed to the spatial location of an auditory cue (Spence & Driver, 1996). Such close cross-modal links between audition and vision are clearly adaptively advantageous, as it is often beneficial to pay visual attention to the location of a new sound in the environment. The end result is that the high-resolution fovea is directed to areas of interest, allowing detailed processing of visual stimuli of interest. Sometimes, though, we will be aware that a nearby sound will be particularly important to pay attention to, such as the horn of a vehicle, or that a sound is not especially relevant, such as the sound of somebody sneezing.

The present study aimed to investigate the extent to which such cross-modal capture of attention might be affected by expectancy. For example, is attention still unavoidably captured by an auditory event if it is unlikely that an interesting visual stimulus will appear in that location (i.e., when it is more strategically beneficial to direct attention to an alternative location)? To our knowledge, this question has not yet been addressed in the cross-modal attention literature. Conversely, will there be any additional attentional effects, over and above attention capture, if the location of an auditory stimulus is particularly likely to share the location of a relevant visual event? Schmitt *et al.* (2000) investigated this latter question, by comparing the visual attention-capturing properties of an auditory cue with 50% versus 80% likelihood of ipsilateral target appearance. These authors found that the auditory cue with 80% validity elicited greater orienting effects only under conditions of greater response complexity; specifically, when participants responded to the lateralised visual target with a left/right localisation or up/down discrimination response, but not when a single button press detection response was required. However, in a study with four potential target locations, Schmitt *et al.* (2001) found either no effect or a 'negligible' effect of increasing auditory cue validity – from uninformative to 80% likelihood of nearby target appearance – both with visual target detection and spatial localisation tasks. We also explore the effects of different response methods, as will be described below.

This study is concerned with both exogenous and endogenous attentional orienting processes, and the manner in which these mechanisms interact in a cross-modal situation. Under 'neutral' conditions – when the target may appear on either the cued or uncued side of the display with equal probability – exogenous orienting can be observed in isolation. When the cue stimulus conveys information concerning the likely side of the impending target, we anticipate that attention will not only be captured in an exogenous fashion by the peripheral cue, but that some attempt will be made to endogenously direct attention to the expected target location. Sometimes this expected location may coincide with the location from which the sound was emitted, but sometimes the expected location may differ from the source of the signal. Here

DOI 10.16910/jemr.2.3.4
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

we evaluate the extent to which these two mechanisms are able to compete with – or complement – each other, depending on the different expectancy conditions.

In this study we also sought to examine the effects of cross-modal attentional orientation on two different dependent measures: saccadic eye movement latencies to the target location and spatial discrimination of target localisation. We used saccadic eye movement latencies as the primary dependent measure in Experiment 1. Although warning effects of auditory stimuli on eye movements to visual stimuli have been studied – e.g., Frens *et al.* (1995) showed that saccades to a visual stimulus are faster when irrelevant auditory stimuli occur in close spatial and temporal alignment – this method of response has not previously been investigated in cross-modal cueing studies concerned with endogenous orienting.

A distinction is usually drawn in the visual attention literature between overt versus covert attention shifts. An overt attention shift occurs when the eyes, head, or entire body move to align the fovea with a new object of interest. While the focus of attention may, in this way, coincide with the area of the visual field to which the fovea is directed, the two are also potentially dissociable. A covert shift of visual attention occurs when the focus of attention is directed to an area of the peripheral or parafoveal visual field independent of any overt movements (Posner, 1980).

While drawing a distinction between overt and covert attention may be descriptively useful, it is worth bearing in mind the strong links between these processes (e.g., Findlay & Gilchrist, 2001), particularly when overt attention is defined in terms of the direction of gaze. For example, evidence from a number of influential studies suggests that, prior to any eye movement, covert visual attention must first be focused on the destination of the saccade (e.g., Shepherd *et al.,* 1986; Kowler *et al.*, 1995; Duebel & Schneider, 1996; although see Stelmach *et al.* (1997) who suggest that this may not apply in some instances of endogenous attentional orientation). Similarly, the premotor theory of attention (Rizzolatti *et al.*, 1987; Sheliga *et al.*, 1995) maintains that a covert attention shift to a given location is equivalent to a programmed, but unexecuted saccade to that location. Given that overt and covert visual attention are intimately related, saccadic reaction time may be regarded as a convenient tool to investigate the triggering of covert attention. Although overt responses are being measured, the observed latencies provide a highly valid index of the time taken to orient covert attention.

In Experiment 2, we used a discrimination measure as the dependent variable, whereby participants needed to discriminate the spatial localisation of a peripherally presented visual target while central fixation was maintained. As some researchers have argued (e.g., Spence & Driver, 1997), this measure of attention is a less contaminated measure of covert attention, since this measure eliminates the confounding influence of 'response priming' by the cue, and the problem of a shift in target detection criterion to the cued location. Experiment 2 used the same auditory cue as in Experiment 1. If the cue serves to attract covert attention in this situation, then this will add support to the assumption that any saccadic latency advantages in Experiment 1 reflect covert attention-attracting properties of the cue.

Experiment 1

The experiments reported here used a methodology analogous to Posner's (1980) spatial cueing task. In essence, this task involves the presentation of a cue stimulus followed by a target, to which a response is required. This study differs from the standard spatial cueing task insofar as a lateral auditory cue, rather than a visual cue, precedes the onset of the visual target stimulus. The auditory stimuli used in this study were presented in free-field form, rather than over headphones, allowing the auditory and visual target stimuli to either be in very close proximity to each other, or to be emitted from different spatial locations. In Experiment 1, in order to measure the relative speed of detection of the visual target stimulus, we measured latency to initiate a saccadic eye movement to the target, which is assumed to provide a measure of the focus of attention.

One issue that this experiment was designed to address was whether reflexive capture of visual attention would occur in response to an auditory cue stimulus. To address this issue, a 'neutral' condition was included, in which the spatial location of the auditory cue was uninformative with respect to the location of the impending visual target. In addition to any potential effect of visual attention capture by the auditory cue, we were also interested in the extent to which expectancy – or cue informativeness – might influence such attentional orientation. To address this issue, two conditions were included in

DOI 10.16910/jemr.2.3.4
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

which the contingent relationship between the cue and target locations was manipulated. In one condition, participants were aware that the target was more likely to appear near the cued location, and in the other condition participants were aware that the target was more likely to appear in the alternative location (i.e., in the opposite visual hemifield from the cue's location).

Three specific questions, then, were under investigation in this study. (1) To what extent does visual attention reflexively shift to the spatial location of the auditory stimulus? (2) Might informativeness confer an additional advantage to the cued location – over and above the uninformative condition – when the target is expected to occur at the cued location? (3) Can cue informativeness lead to rapid shifts of attention to the *uncued* location, when the target is unlikely to occur at the cued location?

We used a 200 ms stimulus onset asynchrony (SOA) between the onset of the auditory cue and the onset of the visual target. Although this is longer than is thought to be optimal for eliciting attention capture by a visual cue stimulus, this SOA has previously been shown to be conducive to the capture of visual attention by a peripheral auditory cue (McDonald *et al.*, 2000; Schmitt *et al.*, 2000; Spence & Driver, 1997).

Methods

Participants

A total of 12 volunteers (7 women), with a mean age of 20.8 years, participated. All participants had normal or corrected-to-normal vision, and reported normal hearing. All participants were paid for participation, and all gave written informed consent. The study was approved by the Lancaster University Psychology Department ethics committee.

Stimuli

The visual stimuli were LEDs, and the central fixation LED was red. There were six green peripheral LEDs, three on the left hand side of the display and three on the right. On each side of the display, one of the three green LEDs (the target stimulus) was located on the same horizontal plane as, but 17.5º more peripheral than, the central red LED. The remaining two green LEDs on each side of the display were positioned 0.5º above and below the target LEDs. These four green peripheral LEDs functioned as place-markers, essentially indicating by flanking the two potential locations of the target stimuli (see Figure 1). Auditory cues were generated by two buzzers presented in free-field form, with one positioned on the left of the display and one on the right, 0.5º more eccentric than the visual targets. Each buzzer emitted a pure tone with a frequency of 2.3 kHz, presented at a sound level of 75 dB (SPL) measured from participants' position. Subjectively, the auditory cues were clearly audible, and clearly localisable with respect to the side from which they were emitted.

Procedure

Participants were seated 168 cm from the display. The display comprised a board, measuring 122 cm (width) x 42 cm, on which the LEDs were positioned and the buzzers were affixed. Participants undertook the experiment in a darkened, sound attenuated room, while the experimenter was located in the adjacent room. Latency to initiate a saccadic eye movement to the visual target was measured, using a Skalar IRIS eye monitoring system with a 500 Hz sampling rate. Only movements of the right eye were analysed. Correct responses were classed as the first saccade (>1º from fixation) initiated towards the visual target between 80 ms and 700 ms after its onset. Trials on which participants were not fixating centrally (within 1º of fixation) on presentation of the visual target were discarded.

There were three conditions, differing with respect to the likelihood of the visual target appearing near the location from which the auditory cue had sounded. In a 50% target-at-cue (TAC) condition, the target appeared on either the cued or uncued side of the display with equal probability. The target appeared in close proximity to the spatial location of the auditory cue on average on four out of every five cued trials in an 80% TAC condition, and on one out of every five cued trials in a 20% TAC condition.

The three experimental conditions were presented in separate blocks. Each block consisted of 88 trials, and was divided into two sub-blocks of 44 trials to allow recalibration if necessary at the halfway point, and to afford participants a brief rest. Order of exposure to the three conditions was counterbalanced. Before each block of trials, participants were informed of the contingent relationship between the cue and target's spatial locations.

DOI 10.16910/jemr.2.3.4
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

Catch trials, where an auditory cue sounded but no visual target appeared, were randomly interspersed throughout the experiment, with eight catch trials presented in each block of 88 trials. At the beginning of the experiment, participants were warned about the occurrence of catch trials, and were requested to attempt to withhold an eye movement on such trials. A catch trial error was defined as a catch trial on which an eye movement (>1° from fixation) was made within 700 ms after cue onset.

Trial Sequence

On a given trial, there was a 1500 ms intertrial interval (ITI) during which the display was blank. A trial began with a red LED presented at a central fixation point, and two green 'place-marker' LEDs on each side of the display. This 'place-marker' display was presented for an interval that varied randomly from 500 to 1500 ms. An auditory cue then sounded for 100 ms, while the 'place-marker' display remained visible. The cue was emitted from either the left or the right side of the display, at a position slightly (0.5°) more eccentric than the spatial location of the targets' potential locations. With the 'place-marker' display still present, a further 100 ms elapsed following the offset of the auditory cue, yielding an SOA of 200 ms. At this time, provided the trial was not a catch trial, a target stimulus – a green LED – was presented. The target was presented, for 800 ms, on the horizontal meridian 17.5° from central fixation, either on the left or the right, midway between the two vertical flankers on the respective side of the display. (On catch trials, the display simply remained as it was for this 800 ms period). Following the offset of the target, the display reverted to the 'place-marker' format, and remained as such for a further 500 ms. See Figure 1 for a diagrammatic representation of one potential trial type.

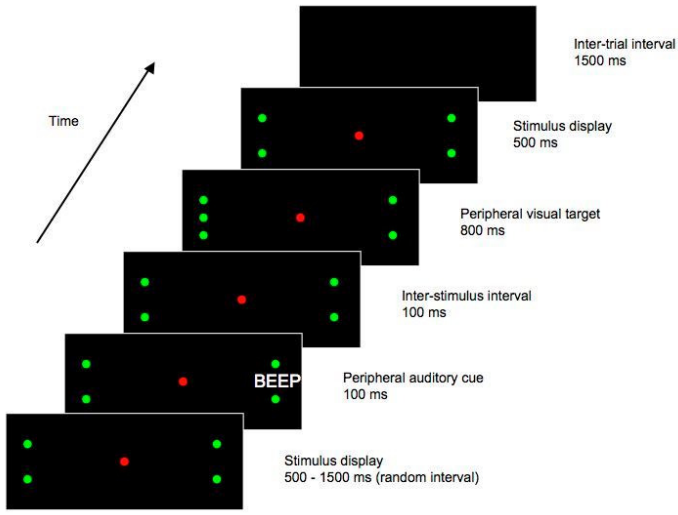


Figure 1: Stimulus sequence from Experiment 1. This figure shows one possible trial type, in which the target appears contralateral to the cue. In this example, the auditory cue is emitted from the right hand side of the display, and the subsequent target is presented on the left. Both cues and targets could also be presented on the opposite side of the display to that shown here, and on catch trials a target did not appear. Figure not to scale.

Results

Saccadic Latencies

Only saccadic latencies greater than 80 ms and less than 700 ms were included in the analysis, thereby excluding 7.7% of trials, and results were collapsed across target side (i.e., left versus right). Using SPSS (v.11), latencies of saccades in the correct direction (i.e., towards the target) were entered into an ANOVA with two within-subject repeated measures factors: Expectancy (20%, 50%, and 80% TAC) and Validity (target ipsilateral to the cue versus target contralateral to the cue).

The ANOVA revealed a significant main effect of Validity ($F(1,55) = 67.78$, $MSE = 778.4$, $p < .001$). Saccades to targets ipsilateral to the auditory cue (266 ms) were faster than saccades to targets contralateral to the cue (320 ms) – see Figure 2. The main effect of Expectancy did not approach significance ($F(2,55) = 0.396$, $p = 0.675$). The ANOVA did, however, reveal a significant interaction between Expectancy and Validity ($F(2,55) = 3.56$, $MSE = 778.4$, $p = .035$), indicating that the effect of validity was not equivalent across the three expectancy conditions.

DOI 10.16910/jemr.2.3.4
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

In order to evaluate the hypotheses concerning the potential differential effects that Expectancy might have on the validity effect, two *a priori* contrasts were undertaken. We compared the size of the observed validity effect in the 50% TAC condition (42 ms) with the validity effect observed in both the 20% (41 ms) and the 80% (79 ms) TAC conditions. The validity effect was reliably larger in the 80% TAC condition compared to the 50% TAC condition ($F(1,55) = 7.122$, MSE = 778.4, $p = .010$). When comparing the 20% and 50% TAC conditions, the size of the validity effect in these two conditions was not reliably different ($F(1,55) = 1.971$, MSE = 778.4, $p = 0.166$).

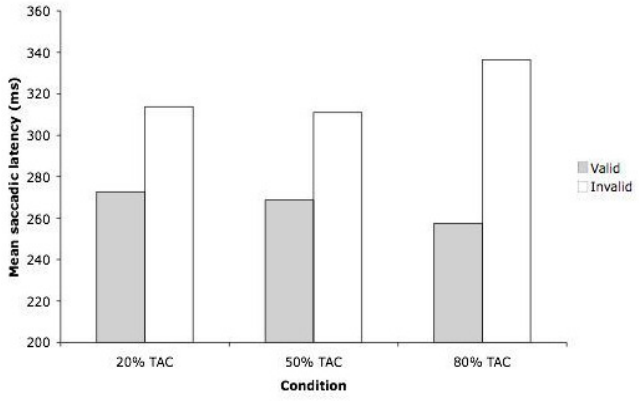


Figure 2: Results from Experiment 1, showing mean saccadic latency (in milliseconds) for valid and invalid trials, across the three conditions. TAC = Target at cue.

Errors

Saccadic eye movements were initiated in the wrong direction (i.e., in the opposite direction from the target's location) on 1.25% of trials in the 20% TAC condition, and on 1.04% of trials in both the 50% and 80% TAC conditions. In each expectancy condition, there were 96 catch trials in total across all 12 participants. In the 20% TAC condition, a total of 17 catch trial errors were made, and 12 of these errors involved participants moving their eyes in the direction of the auditory cue. In the 50% TAC condition, 22 out of a total of 25 catch trial errors involved participants moving their eyes in the direction of the cue. In the 80% TAC condition, 32 out of a total of 33 catch trial errors involved participants moving their eyes in the direction of the cue. These catch trial frequency data were subjected to contingency table analyses. It was found that the number of catch trial errors varied between conditions (Likelihood ratio $\chi^2 = 7.02$, df = 2, $p = .030$), and that there was a trend for catch trial errors to increase from the 20% to the 80% TAC conditions (Linear by linear association $\chi^2 = 6.90$, df = 1, $p = .009$). Furthermore, it was found that the proportion of catch trial errors in the direction of the auditory cue varied between conditions (Likelihood ratio $\chi^2 = 7.13$, df = 2, $p = .028$), with a higher proportion of catch trial errors being made in the direction of the cue as the probability of the target appearing on the cued side increased (Linear by linear association $\chi^2 = 7.03$, df = 1, $p = .008$).

Discussion

A significant validity effect was observed in the 50% TAC condition – participants were faster to launch an eye movement to the target when it occurred near the location of the uninformative auditory cue. One question that this study aimed to address was whether cue informativeness could lead to a cross-modal shift of attention to the uncued location, if the target was unlikely to appear at the cued location. In the 20% TAC condition, cue informativeness clearly did not eliminate the validity effect; indeed, the validity effect was not even reduced relative to the 50% TAC condition. Participants were, in this condition, still faster to look to the target when it appeared on the side of the auditory cue, even though they knew that the target was highly unlikely to appear at that location. We were also interested in whether informativeness would confer an additional advantage to the cued location when the target was also expected (and four times more likely to occur) at the cued location – i.e., in the 80% TAC condition. The results revealed that an additional saccadic latency advantage for the cued location was indeed evident in the 80% TAC condition, over and above that observed in the 50% TAC condition.

What mechanism might underlie the validity effect observed in the three conditions? One possible explanation is that of a criterion shift – a bias to responding to events occurring at the cued location. The pattern of catch trial errors is consistent with this explanation: Participants were more likely to move their eyes in the direction of the auditory cue on catch trials, even in the 20% TAC condition when participants knew that the target was unlikely to appear at the cued spatial location. Furthermore,

DOI 10.16910/jemr.2.3.4
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

the proportion of catch trial errors in which participants moved their eyes in the direction of the cued location increased as the probability of the target appearing on the cued side of the display increased. These data suggest that there was no bias to respond to the *expected* location, but rather to the *cued* location (i.e., the spatial location at which the auditory stimulus had sounded). This explanation is similar to the notion of response priming, or the priming of an ipsilateral response by the cue. As pointed out by Spence and Driver (1997), "quicker responses for targets on the cued side may arise simply because the cue preactivates the appropriate response rather than because of any shift in covert attention." (p.2). If response priming or a criterion shift accounts for the findings in Experiment 1, though, it is interesting that endogenous processes appear to be unable to override the propensity for overt visual attention to be directed to the source of a sound in the environment when it was known that a visual event of interest was unlikely to occur there.

Alternatively, the results of Experiment 1 could be indicative of the orienting of covert visual attention to the cued location. An interpretation of the results in terms of attentional orientation would be as follows: The significant validity effect in the 50% TAC condition suggests that, when the target could appear at either location with equal probability, reflexive attention capture occurred. Attention was captured by the auditory cue, resulting in relatively rapid detection of – and reaction to – the subsequent visual stimulus appearing in the cue's vicinity. When the target appeared in the uncued location, attention needed to be disengaged from the cued location, reoriented to the opposite location, and engaged on this location before a response could be initiated; a time-consuming process. Cue informativeness did not lead to rapid shifts of attention to the uncued but expected location – attention capture was still evident in the 20% TAC condition. The fact that the validity effect was also significant in this 20% TAC condition indicates that the effect of attention capture by the auditory cue is unavoidable or automatic. Even though the target was unlikely to appear at the cued location, an attentional advantage was still evident in this region of the display. The additional advantage for the cued location in the 80% TAC condition – compared to the 50% TAC condition – demonstrates an instance of the exogenous and endogenous attention orienting mechanisms interacting with each other. Not only is visual attention automatically drawn to the location of the sound, but the likelihood of the target appearing nearby leads to attention being deliberately focused on this area of the visual field.

In Experiment 1, we have shown that eye movements to a visual target are faster if the target is presented near the spatial location of an auditory stimulus emitted 200 ms beforehand. But a simple interpretation of this pattern of results from Experiment 1 is challenging. We cannot be completely confident that a shift of visual attention to the cued location underlies the validity effect that was found across all three conditions. It is possible that the results can be partly explained by attention shifting to the cued location, and partly by a bias to respond with a saccadic eye movement to the cued location. Experiment 2 was designed to assess the attention-capturing property of the auditory cue stimulus in the absence of saccadic eye movements. We expected that, by employing these two separate techniques across two experiments, we would provide converging and complementary evidence for the potential cross-modal attention-capturing properties of the auditory stimulus.

Experiment 2

In Experiment 1, the response that was required of participants was an eye movement to the visual target. This raises doubt as to whether any advantage observed at the cued location – in terms of response latencies – is due to attention having shifted to the auditory cue, or if some other, more prosaic explanation may account for the results, such as response priming or a bias to responding to the location of the sound.

Experiment 2 was designed to address this issue by directly assessing the covert attention-capturing property of the same auditory stimulus that was used in Experiment 1. If it is found that the cue stimulus attracts visual attention to its spatial location, then we can be confident that an attentional explanation at least partly underlies the validity effect found in Experiment 1.

In Experiment 2, participants were required to maintain central fixation, while making an elevation judgement (up versus down) of the location of a peripheral visual target presented to the left or to the right. As in the case of Experiment 1, the visual target was presented 200 ms after the onset of an auditory cue stimulus. The cue was emitted from either the left or the right hand side of the display, and the target sometimes appeared on the

DOI 10.16910/jemr.2.3.4
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

same side as the cue, and sometimes in the opposite location. In this case, the response required was orthogonal to the side on which the target appeared – and orthogonal to the direction from which the cue was presented. The cue could not, in this situation, be said to preactivate one of the possible responses. Additionally, any criterion shift – or bias to respond to the cued side of the display over the uncued side – could not facilitate such a discrimination response.

During pilot testing for Experiment 2, stimulus parameters were set such that discrimination judgements were sufficiently demanding to avoid any ceiling effects and therefore response accuracy was the critical dependent variable.

Methods

Participants

The same 12 participants who undertook Experiment 1 also participated in Experiment 2.

Stimuli

The auditory cue stimuli were the same as those used in Experiment 1. The visual stimuli were the same LEDs in the same arrangement as in Experiment 1, although the pattern of their temporal presentation was somewhat different in Experiment 2. An SOA of 200 ms was again used between the onset of the auditory cue and the onset of the visual 'target'.

Procedure

Participants were seated 168 cm from the display board, which was the same as that used for Experiment 1. Eye position was again monitored, and participants were instructed to fixate the red central LED throughout the duration of each trial. As in Experiment 1, three conditions (20%, 50%, and 80% TAC contingencies) were presented in separate blocks of trials. Each block consisted of 80 trials, and each block was divided into two sub-blocks of 40 trials. Order of exposure to the three conditions was counterbalanced. Before each block of trials, participants were informed of the relationship between the cue and target locations. Again, the experiment was conducted in a darkened, sound attenuated room, and the experimenter was in the adjacent room.

Participants were required to indicate, by pressing one of two buttons with any finger, whether they perceived the target stimulus to have appeared above or below the peripheral 'place-marker' LED, regardless of which side it appeared on. Responses were made on a button-box with two buttons; the top button was pressed to indicate 'above' and the bottom button indicated 'below'. Participants were instructed to respond as accurately as possible, and that they should respond before the beginning of the next trial. There were no catch trials in Experiment 2, since such trials are only appropriate to detection responses.

Trial Sequence

On a given trial, there was a 1500 ms ITI during which the display was blank. Next, for a random interval ranging from 500 to 1500 ms, a display with a red central LED and two green LEDs was presented. The green peripheral LEDs were positioned along the horizontal midline of the display, with one to the left and one to the right of the central LED, and were each displaced 17.5º from fixation. These peripheral LEDs functioned as place-markers, indicating the approximate location of the impending target stimulus. An auditory cue then sounded for 100 ms, while the place-marker display remained visible. The cue was emitted from either the left or the right side of the display, at a position slightly (0.5º) more eccentric than the spatial location of the peripheral place-marker LED. With the place-marker display still present, a further 100 ms elapsed following the offset of the auditory cue, yielding (as in Experiment 1) an SOA of 200 ms. At this time, a single green LED – the 'target' stimulus – was presented at one of four potential locations; 0.5º either above or below one of the peripheral green LEDs. After a duration of 100 ms, green LEDs appeared at the remaining three locations (0.5º above and below the peripheral LEDs). This stimulus display, with the central red LED and three green LEDs on each side of the display, remained for 1200 ms, after which the trial ended. See Figure 3 for a diagrammatic representation of one potential trial type.

DOI 10.16910/jemr.2.3.4
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

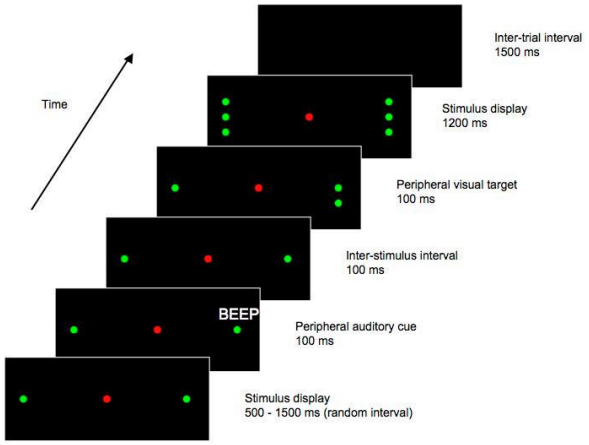


Figure 3: Stimulus sequence from Experiment 2. This figure shows one possible trial type, in which the target appears ipsilateral to the cue. In this example, the auditory cue is emitted from the right hand side of the display, and the subsequent target is presented on the right. Both cues and targets could also be presented on the opposite side of the display to that shown here, and the target could also appear above the central placeholder LED, rather than below as shown here. Figure not to scale.

Results

Any trials with eye movements (>1° from fixation within 700 ms of tone onset), or in which participants were not fixating centrally at tone onset, were excluded from the analysis; this led to the exclusion of only 1.98% of trials. In Experiment 2, participants were instructed to respond as accurately as possible, and on each trial could potentially take up to 2800 ms to execute a response. Consequently, response latencies were not analysed – rather, we analysed the proportion of correct responses. The proportion of correct responses was calculated (out of all possible responses), for each condition (20%, 50%, and 80% TAC conditions) and trial type (target on same side as cue and target on opposite side to cue); these proportions are shown in Table 1. Using SPSS (v.11), the data were subjected to binary logistic regression analysis, with two factors: Expectancy and Condition.

The analysis revealed a significant Validity effect (χ^2 = 6.09, df = 1, p = .014). The proportion of correct responses was significantly higher for targets appearing on the same side as the auditory cue (0.7) versus targets appearing on the opposite side to the auditory cue (0.66). The analysis also revealed a significant effect of Expectancy (χ^2 = 9.56, df = 2, p < .01). *Post-hoc* analysis revealed that overall performance was enhanced in the 50% TAC condition (0.71) compared to both the 20% (0.67) and 80% (0.65) TAC conditions, which did not differ. No significant interaction was found between Expectancy and Validity ($\chi^2 < 1$).

Table 1
Results of Experiment 2, showing proportion correct (out of all possible responses) across the three expectancy conditions.

	Expectancy condition		
Target side	20% TAC	50% TAC	80% TAC
Same side as cue	0.703	0.725	0.672
Opposite cue	0.645	0.688	0.635

Note. *TAC = target at cue.*

Discussion

Discrimination performance was significantly better for the side from which the auditory cue had sounded. This overall outcome mirrors that of Experiment 1, in which we found an overall saccadic latency advantage for the cued location relative to the uncued location. The design of Experiment 2, whereby the potential target location was orthogonal to the direction of the auditory cue, rules out an explanation based on bias or response priming. We can, therefore, be more confident that an attentional effect accounts for the perceptual advantage observed at the cued location.

In the present experiment, the difference between performance accuracy on same side versus opposite trials was relatively small. Indeed, the proportions correct in each expectancy condition, and for same side and opposite side trials, were themselves quite low, in relation to the chance level of 0.5.[1] This undoubtedly reflects the difficulty of the task. If attention was drawn to the appropriate side where the impending target stimulus was about to appear, however, a small but significant perceptual benefit was observed. The results provide evidence for cross-modal attention capture, whereby covert visual attention shifts reflexively to the spatial location of the auditory stimulus, facilitating the difficult up/down discrimination. Interestingly, this cross-modal attention cap-

[1] One-sample t-tests revealed that each of the proportions differed significantly from 0.5 (p < .05 in each case).

DOI 10.16910/jemr.2.3.4
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

ture occurred even when the visual target event was unlikely to appear near the auditory cue (i.e., in the 20% TAC condition). This finding was also observed in Experiment 1, and provides further support for the notion that the auditory cue captures attention in an unavoidable sense. Although it was more strategically beneficial to orient attention to the uncued location, participants were again unable to do so. In both experiments, then, it appears that the bottom-up effect of the auditory cue stimulus overrides any consciously directed attempts to shift visual attention away from the sound.

The results of Experiment 2 provide some support for an attentional explanation of the eye movement latency findings of Experiment 1, rather than simply ipsilateral response preparation (i.e., the priming of an ipsilateral response by the cue), or simply a criterion shift. The same auditory cue was used in both experiments. Irrespective of the response required, a significant advantage was observed for the cued location relative to the uncued location. The auditory cue does appear to exert an attention-capturing effect on covert visual attention, and it seems reasonable to conclude that this is a significant factor underlying the validity effect observed across the three expectancy conditions in Experiment 1. However, this does not necessarily exclude the possible influence of a supplementary effect of a response bias or criterion shift.

In Experiment 2 there was no evidence for differential effects of attentional orientation in the three different expectancy conditions. Orienting was just as successful in the 20% TAC condition as it was in the 80% TAC condition. This finding differs from the outcome of Experiment 1, in which the 80% TAC condition yielded a validity effect of greater magnitude than that of the 50% TAC condition. It would appear that, as Schmitt *et al.* (2000) also found, the characteristics of the target detection task can affect the pattern of results obtained. However, Schmitt *et al.* found that greater response complexity yielded larger orienting effects in their 80% validity condition, whereas in our study we see the opposite; with greater response complexity (Experiment 2), there is a reduced orienting effect in the 80% TAC condition compared to Experiment 1. The differential pattern of results may reflect a factor in relation to the response mode – for example, a greater validity effect might be observed with eye movement responses due to a bias to respond to the cued side when eye movement responses are required. This possibility will be considered further below.

The finding, in Experiment 2, of a better discrimination performance in the 50% TAC condition was unexpected. It is important to note that this finding does not reflect any differential effect of validity across the three expectancy conditions. The size of the validity effect was equivalent across the three expectancy conditions. It would appear that under these conditions, an expectation of likely target side has a small negative effect on discrimination performance, relative to a situation where the target location is uncertain.

Conclusions

In normal individuals, visual attention is reflexively drawn to the spatial location of an auditory stimulus. In the neutral (i.e., 50% TAC) conditions of both experiments reported here we see evidence for such cross-modal attention capture by the auditory stimulus, whereby processing advantages were observed at its spatial location (see also Spence & Driver, 1997; Schmitt *et al.*, 2000; Schmitt *et al.*, 2001; Mazza *et al.,* 2007). Furthermore, it appears that visual attention is *automatically* attracted to the spatial location of a sound, since processing advantages were observed for the location of our auditory cue even when participants were aware that the visual event of interest was most likely to appear elsewhere - a finding that has not been shown before. These results have implications for the existence of "hardwired, structural links between audition and vision in the control of covert attention" (Spence & Driver, 1997). Such rapid cross-modal attention capture is clearly an adaptively useful feature of the human attentional orientation mechanism, allowing us to quickly inspect an area of the visual field from which a sudden noise occurs. Our findings also demonstrate that eye movement latencies to a distant peripheral target location are reduced if an auditory cue sounds in the near vicinity shortly before the target's appearance.

Exactly the same cue was used in both of the experiments reported here. The findings from Experiment 2 suggest that, once visual attention was reflexively captured by this auditory cue, visual perception was improved in its vicinity. Consequently, fine localisation discrimination was superior in the spatial location from which the cue had sounded, relative to the alternative location. In Experiment 1, participants responded with saccadic eye movements more rapidly to a visual target

DOI 10.16910/jemr.2.3.4
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

presented in close proximity to the spatial location of the sound, relative to the opposite location. One interpretation of this finding is that a visual stimulus appearing at the cue's location was detected more rapidly in peripheral vision, due to an exogenous shift of covert attention. Another possibility, though, is that the auditory cue merely primed the saccadic response in its direction.

This idea of response priming, or ipsilateral response preparation, is problematic for numerous spatial cueing experiments – both unimodal and cross-modal. Experiments that require target detection responses, such as eye movements to the target or left/right key-presses, fail to distinguish between attention having shifted to the location of the cue and a response bias towards the cued location. In spite of this, numerous spatial cueing experiments with target detection responses (finding validity effects) have claimed to have shown attentional effects (e.g., Posner, 1980; Corbetta *et al.*, 1993; Abrams & Dobkin, 1994; Sheliga *et al.*. 1995; Rosen *et al.*, 1999; Briand *et al.*, 2000; Lambert *et al.*, 2000). Are we then required to reinterpret a vast range of studies, including many spatial attention cueing studies with both eye movement and manual responses?

One alternative possibility is to reconsider the concept of ipsilateral response preparation. As mentioned above, overt and covert attention are strongly linked. For instance, covert attention shifts in advance of a saccadic eye movement to its landing point (e.g., Shepherd *et al.,* 1986; Duebel & Schneider, 1996; Kowler *et al.*, 1995). Also, according to the premotor theory of attention (Rizzolatti *et al.*, 1987; Sheliga *et al.*, 1995) a covert visual attention shift *is* essentially a planned and programmed, but unexecuted, eye movement. Ultimately, oculomotor programming may be indistinguishable from the orienting of covert visual attention. According to this conceptualisation, the results of Experiment 1 are interpretable in terms of covert attention. Eye movements to the cued location would be difficult to avoid *because* attention had oriented to that location on detection of the auditory cue. However, since an identical cue was used in both experiments, a reasonable assumption is that attention would be attracted to a similar degree in both experimental conditions. Yet we observed a difference between the pattern of findings across the two experiments: In Experiment 1, the advantage conferred by the cue was differentially affected by the expectancy conditions, whereas in Experiment 2 the advantage for the cued location, in terms of discrimination of the target's location, did not differ across the three conditions.

One interpretation that could integrate these data is that both experiments reveal the effects of attention capture by the auditory cue, but that in Experiment 1 the results are also influenced by a bias for moving the eyes to the cued location. This bias to respond to the cued side would explain the larger validity effect of the 80% TAC condition in Experiment 1. A bias to respond to the cued side of the display would also explain the pattern of catch trial errors, whereby the number of such errors, as well as the proportion of catch trial errors in the direction of the cue, increased with increasing expectancy of ipsilateral target appearance across the three conditions.

The experiments reported here suggest that attention is a significant component of validity effects in cross-modal cueing tasks. Both of the measures used here revealed a significant advantage for the location of the auditory cue, in terms of eye movement latencies following target detection and fine spatial discrimination of target localisation. Furthermore, orienting visual attention to a sudden sound appears to be unavoidable, since it occurs even when the source of the sound is not expected to coincide with a target visual stimulus. The influence of the endogenous orienting system was, at the SOA employed here, unable to override exogenous orienting to the location of the auditory cue in the 20% TAC conditions of both experiments. Not only did this unavoidable cross-modal attention capture facilitate the speed of eye movements to align the fovea with the location from which the sound was emitted, but it also served to enhance the processing of visual information at that location in the absence of eye movements.

Acknowledgements

This work was funded by the New Zealand Foundation for Research, Science & Technology (grant no. LANC0401). We are grateful to two reviewers (Robin Walker and Valerie Benson) for helpful comments on an earlier draft of this article.

DOI 10.16910/jemr.2.3.4
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

References

Abrams, R. A., & Dobkin, R. S. (1994). Inhibition of return: Effects of attentional cuing on eye movement latencies. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 467-477.

Bonnel, A. M., Possamaï, C. A., & Schmitt, M. (1987). Early modulation of visual input: A study of attentional strategies. *Quarterly Journal of Experimental Psychology, 39A,* 757-776.

Briand, K. A., Larrison, A. L., & Sereno, A. B. (2000). Inhibition of return in manual and saccadic response systems. *Perception and Psychophysics, 62,* 1512-1524.

Buchtel, H. A., & Butter, C. M. (1988). Spatial attention shifts: Implications for the role of polysensory mechanisms. *Neuropsychologia, 26,* 499-509.

Cheal, M., & Lyon, D. R. (1991). Central and peripheral precuing of forced-choice discrimination. *Quarterly Journal of Experimental Psychology, 43A,* 859-880.

Corbetta, M., Miezin, F. M., Shulman, G. L., & Petersen, S. E. (1993). A PET study of visuospatial attention. *The Journal of Neuroscience, 13,* 1202-1226.

Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research, 36,* 1827-1837.

Findlay, J. M., & Gilchrist, I. D. (2001). Visual attention: The active vision perspective. In L. R. Harris and M. Jenkin (Eds.), *Vision and Attention* (pp. 83-103). Berlin: Springer Verlag.

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 1030-1044.

Frens, M. A., Van Opstal, A. J., & Van der Willigen, R. F. (1995). Spatial and temporal factors determine auditory-visual interactions in human saccadic eye movements. *Perception and Psychophysics, 57,* 802-816.

Jonides, J. (1981). Voluntary versus automatic control over the mind's eye's movement. In J. B. Long and A. D. Baddeley (Eds.), *Attention and Performance IX* (pp. 187-203). Hillsdale, NJ: Erlbaum.

Kowler, E., Anderson, E., Dosher, B., & Blaser, E. (1995). The role of attention in the programming of saccades. *Vision Research, 35 (13),* 1897-1916.

Lambert, A., Norris, A., Naikar, N., & Aitken, V. (2000). Effects of informative peripheral cues on eye movements: Revisiting William James' "derived attention". *Visual Cognition, 7,* 545-569.

Mazza, V., Turatto, M., Rossi, M., & Umiltá, C. (2007). How automatic are audiovisual links in exogenous spatial attention? *Neuropsychologia, 45,* 514-522.

McDonald, J. J., Teder-Sälejärvi, W. A., & Hillyard, S. A. (2000). Involuntary orienting to sound improves visual perception. *Nature, 407,* 906-908.

Müller, H. J., & Rabbitt, P. M. A. (1989). Reflexive and voluntary orienting of visual attention: Time course of activation and resistance to interruption. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 315-330.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32,* 3-25.

Rizzolatti, G., Riggio, L., Dascola, I., & Umiltá, C. (1987). Reorienting attention across the horizontal and vertical meridians: Evidence in favor of a premotor theory of attention. *Neuropsychologia, 25,* 31-40.

Robertson, L. C., & Rafal, R. (2000). Disorders of visual attention. In M. S. Gazzaniga (Ed.), *The New Cognitive Neurosciences* (pp. 633-649). Cambridge, MA: MIT Press.

Robinson, D. L. & Kertzman, C. (1995). Covert orienting of attention in macaques: III. Contributions of the superior colliculus. *Journal of Neurophysiology, 74,* 713-721.

Rosen, A. C., Rao, S. M., Caffarra, P., Scaglioni, A., Bobholz, J. A., Woodley, S. J., Hammeke, T. A., Cunningham, J. M., Prieto, T. E., & Binder, J. R. (1999). Neural basis of endogenous and exogenous spatial orienting: A functional MRI study. *Journal of Cognitive Neuroscience, 11,* 135-152.

Schmitt, M., Postma, A., & De Haan, E. (2000). Interactions between exogenous auditory and visual spatial attention. *Quarterly Journal of Experimental Psychology, 53A,* 105-130.

DOI 10.16910/jemr.2.3.4

This article is licensed under a
Creative Commons Attribution 4.0 International license.

ISSN 1995-8692

Schmitt, M., Postma, A., & De Haan, E. (2001). Cross-modal exogenous attention and distance effects in vision and hearing. *European Journal of Cognitive Psychology, 13,* 343-368.

Sheliga, B. M., Riggio, L., & Rizzolatti, G. (1995). Spatial attention and eye movements. *Experimental Brain Research, 105,* 261-275.

Shepherd, M., Findlay, J. M., & Hockey, R. J. (1986). The relationship between eye movements and spatial attention. *Quarterly Journal of Experimental Psychology, 38A,* 475-491.

Spence, C., & Driver, J. (1996). Audiovisual links in endogenous covert spatial orienting. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 1005-1030.

Spence, C., & Driver, J. (1997). Audiovisual links in exogenous covert spatial orienting. *Perception and Psychophysics, 59,* 1-22.

Spence, C. J., & Driver, J. (1994). Covert spatial orienting in audition: Exogenous and endogenous mechanisms. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 555-574.

Stelmach, L. B., Campsall, J. M., & Herdman, C. M. (1997). Attentional and ocular movements. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 823-844.

Tepin, M. B., & Dark, V. J. (1992). Do abrupt-onset peripheral cues attract attention automatically? *Quarterly Journal of Experimental Psychology, 45A,* 111-132.

Theeuwes, J., Kramer, A. F., Hahn, S., & Irwin, D. E. (1998). Our eyes do not always go where we want them to go: Capture of the eyes by new objects. *Psychological Science, 9,* 379-385.

Wright, R. D., & Ward, L. M. (1994). Shifts of visual attention: An historical and methodological overview. *Canadian Journal of Experimental Psychology, 48,* 151-166.

Walker, R., & McSorley, E. (2008). The Influence of Distractors on Saccade-Target Selection: Saccade Trajectory Effects. *Journal of Eye Movement Research, 2*(3), 1-9.

Yantis, S. (1998). Control of visual attention. In H. Pashler (Ed.), *Attention* (pp. 223-256). Hove, England: Psychology Press / Erlbaum.

Yantis, S., & Jonides, J. (1990). Abrupt visual onsets and selective attention: Voluntary versus automatic allocation. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 121-134.

Zackon, D. H., Casson, E. J., Zafar, A., Stelmach, L., & Racette, L. (1999). The temporal order judgment paradigm: Subcortical attentional contribution under exogenous and endogenous cueing conditions. *Neuropsychologia, 37,* 511-520.

DOI 10.16910/jemr.2.3.4
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692