



# Peripheral visual localization is degraded by globally incongruent auditory-spatial attention cues

Jyrki Ahveninen<sup>1</sup> · Grace Ingalls<sup>2</sup> · Funda Yildirim<sup>2</sup> · Finnegan J. Calabro<sup>2,4</sup> · Lucia M. Vaina<sup>1,2,3</sup>

Received: 20 January 2019 / Accepted: 7 June 2019 / Published online: 14 June 2019  
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

## Abstract

Global auditory-spatial orienting cues help the detection of weak visual stimuli, but it is not clear whether crossmodal attention cues also enhance the resolution of visuospatial discrimination. Here, we hypothesized that if anywhere, crossmodal modulations of visual localization should emerge in the periphery where the receptive fields are large. Subjects were presented with trials where a Visual Target, defined by a cluster of low-luminance dots, was shown for 220 ms at 25°–35° eccentricity in either the left or right hemifield. The Visual Target was either Uncued or it was presented 250 ms after a crossmodal Auditory Cue that was simulated either from the same or the opposite hemifield than the Visual Target location. After a whole-screen visual mask displayed for 800 ms, a pair of vertical Reference Bars was presented ipsilateral to the Visual Target. In a two-alternative forced choice task, subjects were asked to determine which of these two bars was closer to the center of the Visual Target. When the Auditory Cue and Visual Target were hemispatially incongruent, the speed and accuracy of visual localization performance was significantly impaired. However, hemispatially congruent Auditory Cues did not improve the localization of Visual Targets when compared to the Uncued condition. Further analyses suggested that the crossmodal Auditory Cues decreased the sensitivity ( $d'$ ) of the Visual Target localization without affecting post-perceptual decision biases. Our results suggest that in the visual periphery, the detrimental effect of hemispatially incongruent Auditory Cues is far greater than the benefit produced by hemispatially congruent cues. Our working hypothesis for future studies is that auditory-spatial attention cues suppress irrelevant visual locations in a global fashion, without modulating the local visual precision at relevant sites.

**Keywords** Attention · Auditory · Crossmodal · Spatial · Visual

## Introduction

Several previous studies support the view that auditory-spatial orienting cues benefit visual detection and search performance (Perrott et al. 1991; Spence and Driver 1997; Bolia et al. 1999; McDonald et al. 2000; Hanada et al. 2019). In a classic crossmodal attention study, visual elevation judgments were faster and more accurate when an uninformative Auditory Cue was presented on the same rather than the opposite side of the Visual Target (Spence and Driver 1997). These effects were subsequently shown to be related to crossmodal modulation of perceptual sensitivity, instead of post-perceptual biases, i.e., “response selection” (McDonald et al. 2000). Subsequent studies have verified that the spatial congruency of crossmodal Auditory Cues is important for visual detection (Yang and Yeh 2014). However, many of these previous studies utilized comparisons of valid vs. invalid cues: Although this is a powerful way to demonstrate

---

✉ Jyrki Ahveninen  
jyrki@nmr.mgh.harvard.edu

<sup>1</sup> Harvard Medical School, Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital, Charlestown, MA, USA

<sup>2</sup> Brain and Vision Research Laboratory, Department of Biomedical Engineering, Boston University, Boston, MA, USA

<sup>3</sup> Department of Neurology, Harvard Medical School, Massachusetts General Hospital and Brigham and Women's Hospital, Boston, MA, USA

<sup>4</sup> Department of Psychiatry and Bioengineering, University of Pittsburgh, Pittsburgh, PA, USA

more global crossmodal influences, it does not address the question of whether auditory-spatial orienting cues enhance the spatial accuracy of visual processing, as is the case with intramodal visuospatial cues (Yeshurun and Carrasco 1999, 2000).

A well-documented property of multisensory interactions is that they become more significant when the unimodal aspects are weak or noisy (Stein and Stanford 2008). Since receptive fields are smallest in foveal vision, visual information dominates multisensory spatial perception in the central field (Colavita 1974; Hairston et al. 2003). However, a stimulus displayed peripherally is more difficult to detect. Whereas in the central field visual angle differences detectable to humans are tens of times smaller than the corresponding minimum audible angles, in locations around 30° of eccentricity, the relative modality difference is less prominent, corresponding to about 2° vs. 5° of minimum detectable separation in the visual vs. auditory domain, respectively (Cowey and Rolls 1974; Colburn 1996). Thus, in this spatial region where the visual ambiguity increases, observers could be expected to be more prone to rely on alternative sensory information, such as auditory-spatial attention cues.

Here, we examined whether Auditory Cues can facilitate localization of Visual Targets in the periphery where visuospatial certainty is lower. In a crossmodal two-alternative forced choice (2AFC) task, Visual Targets were presented either briefly after auditory stimuli, which either served as a cue for the likely target location (i.e., hemispatially congruent Auditory Cue) or as a distractor to direct initial attention to the opposite hemifield (i.e., hemispatially incongruent Auditory Cue), or with no auditory attention cues (i.e., Uncued condition). Our hypothesis was that when the Auditory Cue is in the same visual field with the Visual Target (hemispatially congruent condition), observers' performance is better than when the Auditory Cue and Visual Target are in opposite visual fields (hemispatially incongruent condition).

## Methods

### Observers

Twelve subjects (ten females; 19–23 years) with normal or corrected-to-normal vision, normal visual contrast sensitivity assessed with the Pelli–Robson Contrast Sensitivity Chart (Pelli et al. 1988), and self-reportedly normal hearing ability participated in the study. All procedures were approved by the local human subjects committee at Boston University and all subjects gave informed consent to participate in this research study. All participants were naïve as to the purpose of the experiment. The data of three subjects were excluded from the final data analysis because of their

inability to perform the visual localization task with sufficient accuracy (sensitivity or  $d' = 1$  in the easiest visual localization condition). This resulted in a final sample of nine subjects (eight females).

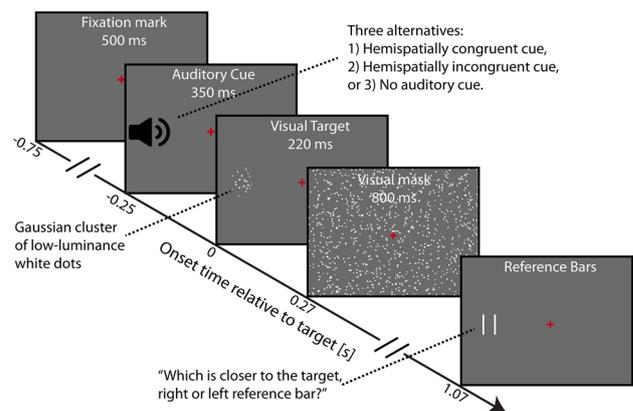
### Experimental setup

The stimuli were generated in BraviShell, a MATLAB-based software package developed in the Brain and Vision Research Laboratory (Biomedical Engineering Department, Boston University, Boston, MA, 2005–2017) built on the Psychophysics Toolbox software package (Brainard, 1997). Visual stimuli were displayed on a 32" Westinghouse monitor with 1440 × 900 resolution. Auditory Cues were presented with Sennheiser HD 280 PRO headphones. Each participant performed the psychophysical tests with the head positioned on a chin rest fixed at 30 cm away from the computer monitor (Fig. 1).

### Stimuli and procedure

#### Visual localization task

The participants were instructed to maintain their gaze on the red crosshair fixation mark (diameter 0.7° of visual angle) displayed at the center of the screen. In each trial,



**Fig. 1** Schematic view of the stimulus and task design. The subjects were asked to fixate on the central cross-fixation mark throughout an entire trial. The Visual Target was a unilateral Gaussian low-luminance dot pattern, which was followed by a whole-screen visual mask and, finally, a pair of vertical Reference Bars that were presented in the same hemifield as the Visual Target but with a slight angular offset. The subjects were asked to report whether the center of the Visual Target was closer to the left or the right Reference Bar. In bimodal trials, the target was preceded by an Auditory Cue, which was either hemispatially congruent or incongruent with the Visual Target. In separate Uncued blocks, the Visual Target was presented without the preceding Auditory Cue. At the end of the trial, the subject was shown a green "O" after correct responses and a red "X" after incorrect responses (not shown in diagram)

the fixation mark appeared 500 ms before the stimulus display and remained on the screen throughout the trial. The Visual Target was a Gaussian cluster of white dots (duration 220 ms, density 2 dots/(°)<sup>2</sup>, radius 5°, dot size 4 min of arc, 1% contrast, background luminance 30.3 Cd/m<sup>2</sup>), presented at random eccentricities between 25° and 35° to the right or left of the central fixation mark. At 50 ms after the offset of Visual Target, the entire screen was covered with a visual mask (duration 800 ms) consisting of white dots (1% contrast, 2 cycles per degree, density 2 dots/(°)<sup>2</sup>; background luminance 30.3 Cd/m<sup>2</sup>). As soon as the mask disappeared, a pair of Reference Bars was displayed. It consisted of two vertically aligned bars, which subtended 10° of visual angle in height and were separated by a gap of 6°. The center of the Reference Bars was situated at 2.5°, 3°, or 3.5° to the left or right of the center of the Visual Target. In a 2AFC task, subjects were instructed to press the keyboard button “1” if the center of the Visual Target had been closer to the left Reference Bar and the keyboard key “2” if it had been closer to the right Reference Bar. A 500-ms visual feedback signal was provided 300 ms after the key press at the same position as the center of the Visual Target. The feedback was either a green “O”, presented after correct responses, or a red “X” displayed after incorrect responses.

### Crossmodal attention cues

The Auditory Cue was a 350 ms broadband noise burst presented 250 ms before the onset of the Visual Target. The source location of Auditory Cue was simulated by adjusting its interaural time difference (ITD; 0.2–0.3 ms for 25°–35° eccentricities) and interaural level difference (ILD; 1.8–2.2 dB for 25°–35° eccentricities). In two-thirds of the cued trials, the source location of the Auditory Cue was congruent with the hemifield of Visual Target. In the other third of the trials, the cue was presented in the hemifield opposite to the target.

### Study visits

During the first of the two visits, subjects were explained the task, signed the informed consent form and were trained on four task blocks of each trial type (24 trials each). In the second visit, the subject performed 2 crossmodal blocks (108 trials each) and 3 blocks with auditory no cues (48 trials each). Each session lasted approximately 40 min. Participants took a short break about every 15 min. Before the beginning of the experiment, participants were adapted to the dark room and the test screen brightness for approximately 300 s. To assure that the observers remembered the task, they were given 4–5 example trials before the beginning of the experimental data collection. The first ten trials of each run were considered as practice and excluded from

the analyses. The experimenter monitored the subject’s gaze fixation on the fixation mark throughout the experiment. All subjects followed the instructions and were able to maintain their gaze direction on the fixation cross throughout each trial.

### Auditory control experiment

We also tested the subjects’ ability to determine the location of the Auditory Cue, per se, relative to the reference bars. The task design was analogous to the main experiment: (1) a 500-ms of fixation was followed by (2) Auditory Cue (for details, see above) and (3) a 475-ms auditory mask (diotic broadband noise; onset 90 ms after the cue), after which (4) the subject was asked to determine which of the subsequently presented Reference Bars was closer to the cue origin. The spatial layout of eccentricities and Reference Bar offsets was identical to the main experiments.

### Behavioral analysis

The accuracy of performance was determined based on the proportion of correct responses ( $P_{\text{Correct}}$ ). Reaction times (RT) were determined from the onset of the Reference Bars in each trial with a correct response. For each subject, we then calculated the median RT for each task condition.

A potential bias in our task was that instead of visuospatial sensitivity, the Auditory Cues could have caused post-perceptual decision (or “response selection”) biases, such as a systematic predisposition to select the bar that is closer (or away from) the contralateral Auditory Cue in the incongruent trials. We, therefore, estimated the subjects’ perceptual sensitivity ( $d'$ ) during the task performance based on Signal Detection Theory (Macmillan and Creelman 1991). This was accomplished by using the  $d'$  function of the psycho package of R (Makowski 2018). The  $d'$  values reflected the subtraction of the  $z$ -transformed  $P_{\text{Correct}}$  and false alarm rates of the responses in our spatial 2AFC task. The *decision bias* (or the “beta value” in the R psycho toolbox nomenclature) was defined as the ratio of the normal density functions at the criterion of the  $z$  values, which were employed in the  $d'$  computation. Here, the *decision bias* value specifically reflected the observer’s predisposition to respond that the target was either closer to the more peripheral or to the more central reference bar, with the unbiased responses having a value of about 1.0.

### Statistical analysis

Linear mixed effects modeling (LMEM) analyses of the task measures were conducted using the `lmer` function of the R `lme4` module (Bates and Maechler 2009; Bates et al. 2015).  $P_{\text{Correct}}$ , RT,  $d'$ , and *decision bias* measures were predicted

using LMEMs, which considered the fixed effects of Cue Type (Uncued, hemispatially congruent Auditory Cue, or incongruent Auditory Cue), the Reference Bar Offset (small, i.e., 2.5°, intermediate, i.e., 3°, or large, i.e., 3.5° difference between the midpoint between the Reference Bars vs. the Visual Target center), as well as the interactions between these predictor terms, and which also controlled for the random effect of subject identity. We then applied a backward stepwise elimination procedure to find the LMEM that best explained the data (“step” function of the lmerTest package). More detailed a priori Helmert contrasts (hemispatially congruent Auditory Cue vs. Uncued; incongruent Auditory Cue vs. the two other conditions combined) were computed from the main LMEM. The degrees of freedom for testing the statistical significance were determined using the lmerTest module of R (main effects and interactions were inferred using the “anova” function, specific a priori contrast using the “summary” function).

## Results

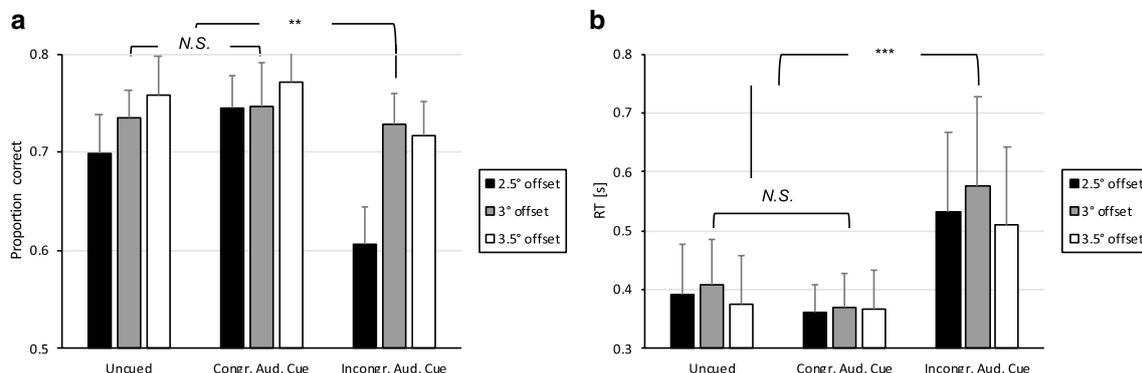
Our LMEMs suggested that whereas hemispatially incongruent Auditory Cues presented at the hemifield opposite to the Visual Target decreased the accuracy, speed, and sensitivity of spatial discrimination of peripheral visual stimuli, hemispatially congruent crossmodal cues produced no improvements in comparison to the Uncued visual-only trials (Fig. 2).

According to the automatic backward-stepwise elimination procedure, the accuracy data were best explained by an LMEM that predicted  $P_{\text{Correct}}$  by the fixed effects of Cue Type and Reference Bar Offset, as well as the random effect of subject identity. This LMEM revealed a significant main effect of Cue Type (Uncued, hemispatially congruent Auditory Cue, or incongruent Auditory Cue;  $F_{2,68}=7.5$ ;

$p < 0.001$ ). The more specific a priori Helmert contrast suggested that this main effect was driven by the significant decrease of  $P_{\text{Correct}}$  in trials with hemispatially incongruent Auditory Cues, compared with the two other conditions ( $t_{68} = -3.9$ ;  $p < 0.001$ ; Fig. 2a). However, the hemispatially congruent Auditory Cues produced no significant improvement of  $P_{\text{Correct}}$  in comparison to the Uncued condition. Finally, the best-fitting LMEM showed that the visual localization performance improved as a function of increasing Reference Bar Offset across all trial types ( $F_{2,68} = 3.8$ ,  $p < 0.05$ ).

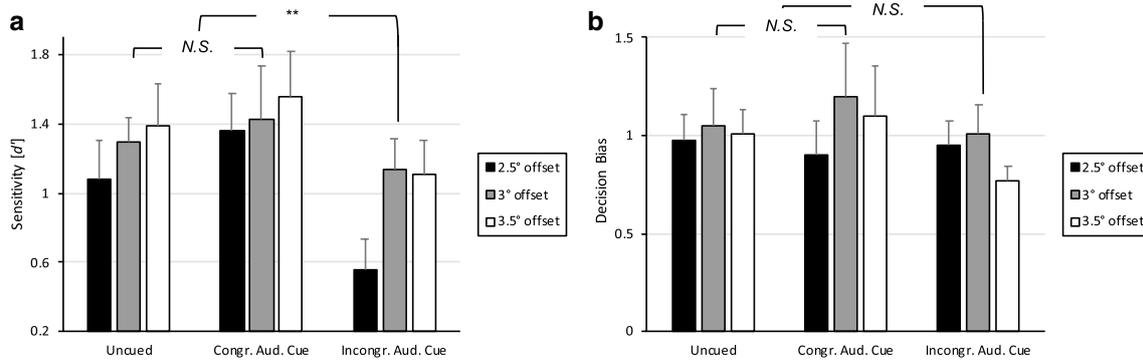
The RT data were best explained by a model that included the fixed effect of Cue Type and random effect of subject identity. Consistent with the  $P_{\text{Correct}}$  results, this LMEM showed a highly significant main effect of Cue Type (Uncued, hemispatially congruent Auditory Cue, or incongruent Auditory Cue;  $F_{2,70} = 11.9$ ,  $p < 0.001$ ). The a priori Helmert contrasts showed that the RTs were significantly delayed in the trials with incongruent Auditory Cues vs. the two other conditions (Helmert contrast,  $t_{70} = 4.8$ ,  $p < 0.001$ ; Fig. 2b), whereas the comparison between trials with congruent Auditory Cues vs. Uncued trials showed no significant differences. The inclusion of terms involving the Reference Bar Offset did not significantly improve the LMEM fit.

The analysis of sensitivity ( $d'$ ) and *decision bias* estimates, computed based on the Signal Detection Theory, were highly consistent with the results of the  $P_{\text{Correct}}$  and RT analyses (Fig. 3). The best-fitting LMEM revealed a significant main effect of Cue Type (Uncued, hemispatially congruent Auditory Cue, or incongruent Auditory Cue;  $F_{2,70} = 5.3$ ,  $p < 0.01$ ; Fig. 3a). The a priori contrasts derived from the main LMEM revealed that  $d'$  was significantly reduced in trials with hemispatially incongruent Auditory Cues, in comparison to the two other conditions (Helmert contrast,  $t_{70} = -3.0$ ,  $p < 0.01$ ), but that there were no differences in  $d'$  between the congruent Auditory Cue vs.



**Fig. 2** Accuracy and speed of target localization in the visual periphery. **a** Group averages of  $P_{\text{Correct}}$ s for different Cue Type and Reference Offset conditions. **b** Group averages of RTs in different Cue

Type and Reference Offset conditions. Error bars refer to the standard error of mean (SEM) across subjects. \*\*\* $p < 0.001$  in Helmert contrasts derived from the main LMEM



**Fig. 3** Sensitivity ( $d'$ ) and *decision bias* values of target localization in the visual periphery. **a** Group averages of  $d'$  estimates in different Cue Type and Reference Offset conditions. **b** Group averages of *decision bias* estimates in different Cue Type and Reference Offset con-

ditions. Error bars refer to the standard error of mean (SEM) across subjects. \*\*\* $p < 0.001$  in Helmert contrasts derived from the main LMEM

Uncued conditions (Fig. 3a). The inclusion of terms involving the Reference Bar Offset did not significantly improve the explanatory power of the LMEM for  $d'$  data.

The LMEM analyses of *decision bias* revealed no significant main effects or interactions (Fig. 3b), consistent with previous studies suggesting that crossmodal attention cues modulate the sensitivity rather than post-perceptual decision making during visual performance (McDonald et al. 2000).

Finally, the auditory control experiment demonstrated that even with the diffuse acoustic backward masking that was not used in the bimodal attention condition, the localization of the Auditory Cue vs. Reference Bars was significantly above chance level (group mean  $P_{\text{Correct}} = 56\%$ ,  $t_8 = 3.6$ ,  $p < 0.01$ ; group mean  $d' = 0.3$ ,  $t_8 = 2.32$ ,  $p < 0.01$ ).

## Discussion

We examined how Auditory Cues presented briefly before the Visual Target stimuli affect spatial localization in visual periphery. The participants were asked to localize the Visual Target in relation to two vertical Reference Bars presented at a slight visual angle offset relative to the center of the Visual Target. In support of our hypotheses, the speed and accuracy of visual localization was significantly better in trials where the Auditory Cue was presented in the same vs. opposite hemifield relative to the Visual Target. Our analyses based on the Signal Detection Theory, further, verified that these effects reflected differences in attentional modulation of perceptual sensitivity rather than decision biases.

Previous studies provide ample evidence for faster RTs, improved accuracy, and increased perceptual sensitivity for visual stimuli when they are preceded by Auditory Cues in the same location (Spence and Driver 1997; McDonald et al. 2000; McDonald and Ward 2000; Schmitt et al. 2000). Based on these findings, a theory has been proposed that

exogenous shifts of spatial attention are governed by a common supramodal system, which can be similarly activated by auditory or visual cues to facilitate the perception of visual stimuli (McDonald et al. 2012). Whereas the statistically significant differences between incongruent vs. congruent Auditory Cue conditions are, in principle, in line with this notion, it is noteworthy that the effects observed in our study appeared to be explained by perceptual degradation after incongruent cues rather than by crossmodal facilitation after congruent cues. That is, in contrast to crossmodally cued orienting to global visuospatial patterns such as motion (Hanada et al. 2019), the present study showed no significant benefits of congruent Auditory Cues in comparisons to the unimodal or Uncued visual condition. This is at odds with the previously reported effect of intramodal visual orienting cues, which help increase the local spatial accuracy of visual perception (Yeshurun and Carrasco 1999, 2000).

A potential explanation for the present pattern of results is that Auditory Cues are not relevant for precise orienting to visual space (e.g., see Fiebelkorn et al. 2011). Whereas fine-grained spatial discrimination is dominated by the visual system (Howard and Templeton 1966; Colavita 1974; Hairston et al. 2003), Auditory Cues could guide attention to sites beyond central vision, routed through the rapid “where” pathway in posterior portions of non-primary auditory cortex (Rauschecker and Tian 2000; Ahveninen et al. 2014). One of the most prominent crossmodal neuronal effects of such Auditory Cues is the increase of alpha oscillations in visual areas representing the irrelevant hemifield (Banerjee et al. 2011; Thorpe et al. 2012; Ahveninen et al. 2013), which is very similar to the effects of hemispacial visual attention cues (Worden et al. 2000) and presumably reflects active suppression of irrelevant locations of visual field (Foxe et al. 1998). Evidence for neuronal effects of Auditory Cues on local visuospatial attention, which would increase the spatial resolution of visual perception the way that intramodal cues

appear to do (Yeshurun and Carrasco 1999, 2000), is much scarcer. Thus, one might speculate that during visual tasks, Auditory Cues activate a broader or global spatial orienting system that triggers larger exogenous attention shifts, and that this helps suppress the representation of irrelevant visual field locations. Such global orienting processes could be based on a modality non-specific or “supramodal” system (McDonald et al. 2012). However, the exogenous shifts of spatial attention do not seem to be able to facilitate more fine-grained, or local, visuospatial attention, which perhaps is controlled by modality-specific information (Ward 1994).

Another possibility is that the occurrence of incongruent auditory orienting cues activates an additional conflict processing system (Botvinick et al. 2004; Huang et al. 2014), which then hinders the speed and accuracy performance at the post-perceptual or “response selection” stage, due to the increased processing demands. However, based on the results of our analysis, the present modulations of accuracy of visual performance by auditory orienting cues reflected changes in perceptual sensitivity rather than decision biases. The notion that present effects cannot be explained by post-perceptual decision biases is also consistent with the predictions of the load theory of visual attention, which suggests that the effect of distractors should decrease as a function of perceptual difficulty (Lavie 2005). In the experiments described here, there was no significant interaction between the cue congruence (or Cue Type) and Reference Bar Offset that reflected the angular offset between the midpoint between the Reference Bars and the center of the Visual Target. Furthermore, in both auditorily cued and Uncued trials, the accuracy and perceptual sensitivity ( $d'$ ) of performance improved as a function of decreasing the Reference Bar Offset.

An important consideration in interpreting our results is the spatial resolution of auditory cueing. The auditory-spatial stimuli were based on a generic ITD/ILD cues, which are not as accurate as individualized simulations including both binaural and monaural spatial cues (for a review, see Ahveninen et al. 2014). Importantly, the spatial differences between the Reference Bars and the Auditory Cues were close to the previously reported discrimination thresholds in the present range of 25°–35° azimuthal directions (Colburn 1996). That is, with the 2.5°, 3°, and 3.5° Reference Bar Offsets, the distance of the further bar (i.e., the incorrect choice) was only 5.5°, 6°, or 6.5° away from the simulated origin of the Auditory Cue. However, in our separate control experiment, the subjects discriminated the locations of Auditory Cues significantly above chance level, even though the cues were followed by a diffuse backward masking sound that probably strongly interfered with the auditory-spatial representation (Ege et al. 2018). In contrast, in the main experiment, the Auditory Cues were not acoustically masked. It is thus reasonable to assume that the auditory attention cueing

was spatially more robust than in the case of the backward-masked control experiment.

Finally, it is necessary to consider whether the timing of stimuli might contribute to the differences in effects observed in the present study and previous studies that measured the effects of auditory-spatial cues on visual detection sensitivity. Here, the Visual Target preceded the onset of Auditory Cue by 250 ms, and the sound cue overlapped the onset of the Visual Target. These timing parameters were selected based on previous studies, which have suggested consistent differences between the effects of valid vs. invalid Auditory Cues even at much shorter cue-target onset asynchronies (e.g., Spence and Driver 1997). A further justification is provided by investigators of unimodal attention, who have argued that at longer cue-to-target onset asynchronies, spatial orienting to auditorily cued locations becomes more increasingly influenced by non-specific alerting as opposed to location-specific effects (Mondor and Zatorre 1995). However, as the majority of previous visual-auditory studies have examined the effects of crossmodal cueing on detection sensitivity, further studies are clearly needed to determine how the timing of crossmodal cues affects the precision of visuospatial attention and discrimination.

In conclusion, hemifield-incongruent auditory-spatial attention cues suppressed irrelevant visual locations, but congruent crossmodal attention cues did not benefit visual precision in comparison to the unimodal condition. Our working hypothesis for future studies, to be tested with more precise auditory-spatial simulations, is that whereas the suppression after hemifield-incongruent Auditory Cues is based on a supramodal system for exogenous shifts of spatial attention, the more fine-grained, local, visual attention is based on a modality-specific system.

**Acknowledgements** This work was supported by the National Science Foundation Grant 1545668 (LMV), and by the National Institutes of Health grants R01DC016765 (JA) and R01DC016915 (JA).

## Compliance with ethical standards

**Conflict of interest** The authors declare that there are no conflicts of interest regarding the publication of this article.

## References

- Ahveninen J, Huang S, Belliveau JW, Chang WT, Hämäläinen M (2013) Dynamic oscillatory processes governing cued orienting and allocation of auditory attention. *J Cogn Neurosci* 25:1926–1943
- Ahveninen J, Kopco N, Jääskeläinen IP (2014) Psychophysics and neuronal bases of sound localization in humans. *Hear Res* 307:86–97
- Banerjee S, Snyder AC, Molholm S, Foxe JJ (2011) Oscillatory alpha-band mechanisms and the deployment of spatial attention to anticipated auditory and visual target locations: supramodal or sensory-specific control mechanisms? *J Neurosci* 31:9923–9932

- Bates DM, Maechler M (2009) lme4: Linear mixed-effects models using S4 classes. In: R package version 0.999999-0
- Bates D, Mächler M, Bolker B, Walker S (2015) Fitting linear mixed-effects models using lme4. *J Stat Softw* 67:48
- Bolia RS, D'Angelo WR, McKinley RL (1999) Aurally aided visual search in three-dimensional space. *Hum Factors* 41:664–669
- Botvinick MM, Cohen JD, Carter CS (2004) Conflict monitoring and anterior cingulate cortex: an update. *Trends Cogn Sci* 8:539–546
- Colavita FB (1974) Human sensory dominance. *Percept Psychophys* 16:409–412
- Colburn HS (1996) Computational models of binaural processing. In: Hawkins H, McMullen T (eds) *Auditory computation*. Springer, New York, pp 332–400
- Cowey A, Rolls ET (1974) Human cortical magnification factor and its relation to visual acuity. *Exp Brain Res* 21:447–454
- Ege R, van Opstal AJ, Bremen P, van Wanrooij MM (2018) Testing the precedence effect in the median plane reveals backward spatial masking of sound. *Sci Rep* 8:8670
- Fiebelkorn IC, Foxe JJ, Butler JS, Molholm S (2011) Auditory facilitation of visual-target detection persists regardless of retinal eccentricity and despite wide audiovisual misalignments. *Exp Brain Res* 213:167–174
- Foxe JJ, Simpson GV, Ahlfors SP (1998) Parieto-occipital approximately 10 Hz activity reflects anticipatory state of visual attention mechanisms. *NeuroReport* 9:3929–3933
- Hairston WD, Wallace MT, Vaughan JW, Stein BE, Norris JL, Schirillo JA (2003) Visual localization ability influences cross-modal bias. *J Cogn Neurosci* 15:20–29
- Hanada GM, Ahveninen J, Calabro F, Yengo-Kahn A, Vaina LM (2019) Cross-modal cue effects in motion processing. *Multisens Res* 32:45–65
- Howard IP, Templeton WB (1966) *Human spatial orientation*. Wiley, London
- Huang S, Rossi S, Hämäläinen M, Ahveninen J (2014) Auditory conflict resolution correlates with medial-lateral frontal theta/alpha phase synchrony. *PLoS One* 9:e110989
- Lavie N (2005) Distracted and confused?: selective attention under load. *Trends Cogn Sci* 9:75–82
- Macmillan NA, Creelman CD (1991) *Detection theory: a user's guide*. Cambridge University Press, Cambridge
- Makowski D (2018) The psycho package: an efficient and publishing-oriented workflow for psychological science. *J Open Source Softw* 3:470
- McDonald JJ, Ward LM (2000) Involuntary listening aids seeing: evidence from human electrophysiology. *Psychol Sci* 11:167–171
- McDonald JJ, Teder-Salejarvi WA, Hillyard SA (2000) Involuntary orienting to sound improves visual perception. *Nature* 407:906–908
- McDonald JJ, Green JJ, Stormer VS, Hillyard SA (2012) Cross-modal spatial cueing of attention influences visual perception
- Mondor TA, Zatorre RJ (1995) Shifting and focusing auditory spatial attention. *J Exp Psychol Hum Percept Perform* 21:387–409
- Pelli DG, Robson JG, Wilkins AJ (1988) The design of a new letter chart for measuring contrast sensitivity. *Clin Vis Sci* 2:187–199
- Perrott DR, Sadralodabai T, Saberi K, Strybel TZ (1991) Aurally aided visual search in the central visual field: effects of visual load and visual enhancement of the target. *Hum Factors* 33:389–400
- Rauschecker JP, Tian B (2000) Mechanisms and streams for processing of “what” and “where” in auditory cortex. *Proc Natl Acad Sci USA* 97:11800–11806
- Schmitt M, Postma A, De Haan E (2000) Interactions between exogenous auditory and visual spatial attention. *Q J Exp Psychol A* 53:105–130
- Spence C, Driver J (1997) Audiovisual links in exogenous covert spatial orienting. *Percept Psychophys* 59:1–22
- Stein BE, Stanford TR (2008) Multisensory integration: current issues from the perspective of the single neuron. *Nat Rev Neurosci* 9:255–266
- Thorpe S, D'Zmura M, Srinivasan R (2012) Lateralization of frequency-specific networks for covert spatial attention to auditory stimuli. *Brain Topogr* 25:39–54
- Ward LM (1994) Supramodal and modality-specific mechanisms for stimulus-driven shifts of auditory and visual attention. *Can J Exp Psychol* 48:242–259
- Worden MS, Foxe JJ, Wang N, Simpson GV (2000) Anticipatory biasing of visuospatial attention indexed by retinotopically specific alpha-band electroencephalography increases over occipital cortex. *J Neurosci* 20:RC63
- Yang YH, Yeh SL (2014) Unmasking the dichoptic mask by sound: spatial congruency matters. *Exp Brain Res* 232:1109–1116
- Yeshurun Y, Carrasco M (1999) Spatial attention improves performance in spatial resolution tasks. *Vis Res* 39:293–306
- Yeshurun Y, Carrasco M (2000) The locus of attentional effects in texture segmentation. *Nat Neurosci* 3:622–627

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.