Guiding spatial attention by multimodal reward cues

Vincent Hoofs[1] · Ivan Grahek[1,2] · C. Nico Boehler[1] · Ruth M. Krebs[1]

Abstract

Our attention is constantly captured and guided by visual and/or auditory inputs. One key contributor to selecting relevant information from the environment is reward prospect. Intriguingly, while both multimodal signal processing and reward effects on attention have been widely studied, research on multimodal reward signals is lacking. Here, we investigated this using a Posner task featuring peripheral cues of different modalities (audiovisual/visual/auditory), reward prospect (reward/no-reward), and cue-target stimulus-onset asynchronies (SOAs 100–1,300 ms). We found that audiovisual and visual reward cues (but not auditory ones) enhanced cue-validity effects, albeit with different time courses (Experiment 1). While the reward-modulated validity effect of visual cues was pronounced at short SOAs, the effect of audiovisual reward cues emerged at longer SOAs. Follow-up experiments exploring the effects of visual (Experiment 2) and auditory (Experiment 3) reward cues in isolation showed that reward modulated performance only in the visual condition. This suggests that the differential effect of visual and auditory reward cues in Experiment 1 is not merely a result of the mixed cue context, but confirms that visual reward cues have a stronger impact on attentional guidance in this paradigm. Taken together, it seems that adding an auditory reward cue to the inherently dominant visual one led to a shift/extension of the validity effect in time – instead of increasing its amplitude. While generally being in line with a multimodal cuing benefit, this specific pattern highlights that different reward signals are not simply combined in a linear fashion but lead to a qualitatively different process.

Keywords Reward · Visual attention · Multimodal cue · Posner cueing task

Introduction

In order to optimize actions and flexibly adapt to changes in the environment, an organism needs to efficiently guide attention towards the most relevant and valuable information (Mobbs et al., 2015). This capacity, which relates to extracting relevant information from increasingly complex contexts, has been a topic of interest in psychological research for many decades (for review, see Driver, 2001). In well-controlled lab environments, the orienting of attention is often studied by embedding motivational valence information in visual attention tasks using reward cues or reward

block manipulations. While some of these studies explored general performance benefits of reward cues when discriminating visual target stimuli at predicted locations (Krebs et al., 2012; Schevernels et al., 2014), others have specifically focused on the effects of reward on orienting (and re-orienting) of attention using valid and invalid spatial cues (Bucker & Theeuwes, 2014; Engelmann & Pessoa, 2007; Milstein & Dorris, 2007; Small et al., 2005). These latter studies rely on the Posner cuing paradigm, hence probing both the initial orienting of attention towards the prioritized (cued) part of the screen, re-centering, and re-orienting towards the opposite part if the target appears in the non-cued location (see Posner, 1980; Wright & Ward, 2008).

For instance, Small et al. (2005) conducted a visual attention task with central spatial cues (i.e., endogenous cues) that could be valid, invalid, or non-directional with regard to the upcoming target location. The focus of the analysis is the so-called validity effect, which reflects the relative performance benefit and cost associated with valid and invalid location cues, respectively. The task was performed under different reward conditions (win/loss/no-reward), manipulated in a

Vincent Hoofs and Ivan Grahek contributed equally to this work.

✉ Vincent Hoofs
vincent.hoofs@ugent.be

1 Department of Experimental Psychology, Ghent University, Henri Dunantlaan 2, 9000 Ghent, Belgium

2 Department of Cognitive, Linguistic & Psychological Sciences, Brown University, Providence, RI, USA

block-wise fashion. They found that reward differentially facilitated target detection at the locations where the events were expected to occur (in valid trials), while loss prospect differentially facilitated responses in invalid trials (albeit both at trend level). A similar but more nuanced pattern was observed in a study employing peripheral incentive cues (i.e., exogenous cues) in a trial-to-trial fashion (Bucker & Theeuwes, 2016). In this study, the specific color of one of two place holders (left/right) predicted win/loss/no-reward prospect prior to the performance of basic target discrimination. It was found that peripheral reward cues induced stronger validity effects compared to no-incentive and loss cues (see also Engelmann & Pessoa, 2007; Munneke et al., 2015). This effect was further modulated by the cue-target interval in that (invalid) reward cues led to longer lingering at the cued location, while loss cues promoted faster disengagement from the cued location (for a related block manipulation, see Bucker & Theeuwes, 2014). This pattern seems to indicate that peripheral reward cues not only lead to stronger attentional capture at the cued location, but also to temporally extended attentional orienting. It is important to note, however, that other related studies observed merely main effects of reward and validity on spatial orienting and no (or no robust) interactions between the two factors (see Baines et al., 2011; Engelmann & Pessoa, 2007). One reason for this might be related to design differences in that the study by Engelmann and Pessoa (2007), for instance, displayed reward cues centrally before the actual spatial cue, which ameliorates a direct influence of reward on perceptual processing at the cued location.

While the before-mentioned studies probed performance-contingent effects of reward on attentional orienting, reward-related stimulus features have also been shown to guide attention in a more incidental fashion in the form of (involuntary) carry-over effects from trial-to-trial or more generally across time (see Anderson, 2016; Camara et al., 2013; Chelazzi et al., 2013; Failing & Theeuwes, 2018; Hickey et al., 2010). For instance, stimulus features that have signaled reward in a previous trial will capture attention in the next trials even when this does not benefit the task goal (Hickey et al., 2010). Similarly, stimulus features that had signaled reward in a pre-task training phase can readily capture attention in the subsequent task phase in which no reward is at stake (Anderson et al., 2011). The effects of these automatic/involuntary (rather than goal-directed/voluntary) reward manipulations are assumed to arise from modulations at early perceptual processing stages (Itthipuripat et al., 2019). Together, the studies discussed so far show that visual incentive cues can modulate performance in visual attention tasks (both via voluntary attentional orienting and involuntary attentional capture), and these manipulations are particularly effective if reward signals directly overlap with target location.

In contrast to the above laboratory studies, real-life situations often feature multimodal information sources – and motivational signals are no exception. Extreme examples include TV commercials and slot machines, but this is also the case when pouring your favorite drink. Despite this, as mentioned above, research into the effects of reward on attention mostly focuses on one modality at a time, and mainly on the visual one. This is not only problematic with regard to ecological validity, but also disregards the notion that many spatial and temporal systems and the associated neural processes are multisensory in nature (Klemen & Chambers, 2012). As such, the present study is inspired by research on reward-modulated attentional guidance on the one hand, and the attentional consequences of multimodal (or multisensory) cuing on the other. The common finding in the multimodal cuing literature is that attentional orienting can be enhanced by presenting multiple concurrent cues of different modalities (visual, auditory, tactile), which has been termed multimodal (cueing) benefit (e.g., Frassinetti et al., 2002; Ho et al., 2009; McDonald et al., 2000; Noesselt et al., 2008; Santangelo, Ho, Spence, 2008; Spence & Santangelo, 2009; Van der Burg et al., 2008; Vroomen & De Gelder, 2000). Of note, other lab studies have failed to find robust multimodal cueing benefits (e.g., Barrett & Krumbholz, 2012; Ngo & Spence, 2010; Santangelo et al., 2006; Spence & Driver, 1999; Ward, 1994). This discrepancy might be related to particular differences in task design, such as task difficulty, in that multimodal cueing benefits were found to be more pronounced under high perceptual load conditions (Santangelo & Spence, 2007; Santangelo, Ho, Spence, 2008). Moreover, it appears that a multimodal cuing benefit is more likely if cues are presented from the same spatial location (e.g., Ho et al., 2009). Despite these mixed results, additive effects of multimodal cues received a lot of attention, also from industry and policy makers to improve efficacy of safety signals in different domains (Haas & Van Erp, 2014; Oskarsson et al., 2012). With regard to the underlying mechanisms of the multimodal cuing benefit, one crucial observation is that sound cues modulate the neurophysiological response of visual cortex to visual targets (McDonald & Ward, 2000), which is indicative of multisensory integration at early sensory-processing areas (for a review, see Driver & Noesselt, 2008). Similar multimodal cuing benefits have been reported for overt attention shifts (i.e., eye-movements) in the form of a linear integration of concurrent visual and auditory signals into a spatial saliency map (Quigley et al., 2007). Such linear integration is assumed to be supported by the superior colliculus, which receives both visual and auditory inputs and plays a crucial role in covert and overt shifts of attention (Meredith & Stein, 1986). Before moving on to the aim of the present study, it is important to note that while multimodal cuing research is strongly overlapping with the concept of multisensory

integration, the latter is not the main focus of the present study (for a discussion of the relationship between multisensory integration and multimodal cuing, see McDonald et al., 2001).

Circling back to the beginning, despite the common interest in reward-induced performance modulations in visual attention tasks on the one hand, and the relevance (and effectiveness) of multimodal signals in daily life on the other, there has been no systematic investigation of the potential added value of multimodal reward signals on attentional orienting to date. To test this, we embedded visual, auditory, and audiovisual reward cues in a visual spatial attention task (Posner, 1980). Our main prediction was that simultaneous presentation of two motivationally relevant signals (i.e., reward prospect cues) would increase attentional guidance in the form of a stronger (additive) validity effect. This pattern, which we label as reward-modulated cuing benefit,[1] would resonate with the linear integration hypothesis put forward for multimodal signals in overt attention research (Quigley et al., 2007). Alternatively, it is possible that the additional reward cue is not further enhancing attentional guidance (because it literally provides the same information), and that the most salient signal guides attention, which corresponds to the maximum hypothesis (see Quigley et al., 2007).

Finally, but least likely, the simultaneous presentation of two salient signals could diminish the validity effects due to an attentional bottleneck (Broadbent, 1958). These predictions were tested in a multimodal cuing experiment (Experiment 1), complemented by two follow-up experiments featuring visual (Experiment 2) and auditory (Experiment 3) cues in isolation.

In keeping with previous related paradigms, we included a manipulation of cue-target interval, operationalized as stimulus-onset asynchrony (SOA[2]), to explore the time course of attentional orienting in response to the different cue types. The range of SOAs was fairly large (100–1,500 ms) to accommodate paradigmatic differences between the different research fields. Specifically, typical Posner tasks with peripheral (exogenous) cues feature SOAs ranging from 100 to 500 ms (Wright & Ward, 2008), while studies exploring the effects of peripheral *reward* cues on attentional

orienting have used longer SOAs ranging from about 150 to 1,000 ms (e.g., Bucker & Theeuwes, 2016). Moreover, studies using central (endogenous) reward cues to predict the upcoming target location featured longer SOAs around 1,500 ms (Krebs et al., 2012; Schevernels et al., 2014). Finally, SOAs in the multimodal cuing literature typically lie between

100 and 700 ms (Spence & Santangelo, 2009). While we did not have specific predictions for all the different combinations of SOAs and other factors due to the hybrid design, we highlight the most relevant observations from the literature here. First, peripheral cues in the Posner task typically yield strong validity effects at around 100 ms after cue onset, which attenuate at longer SOAs (Wright & Ward, 2008). This effect is prolonged for reward-associated peripheral cues (Bucker & Theeuwes, 2016). Second, centrally presented (endogenous) reward cues lead to preparatory attention across the cue-target interval (Schevernels et al., 2014). Third, robust multimodal cuing benefits have been reported for cue-target SOAs of 230–240 ms (Ho et al., 2009; Santangelo & Spence, 2007).

Methods and results

Experiment 1

Methods Experiment 1

Participants All participants in this study were recruited through the local university online recruiting website. These students were between 18 and 35 years of age, right-handed, had normal color and audio perception, as well as normal or corrected-to-normal visual acuity, and no (history of) diagnosed mental disorders (all criteria based on the screening questionnaire embedded in the recruitment website). All three experiments were performed in the laboratory. For the first experiment, data were collected from 44 participants (31 females). The data of two participants were excluded due to a high number of errors (more than two standard deviations from the group mean, averaged across all valid and invalid conditions). The remaining 42 participants (30 females) were on average 22.5 years old (SD = 3.2 years). In addition to a 10 € base payment or one course credit for the 60 min of participation, participants received an average bonus of 3.67 € (about 80% of 5 €, see below).

Paradigm and procedure Participants performed an adapted version of the Posner cueing task (Posner, 1980) in which they could win up to 5 € bonus money (Fig. 1). After a short (no-reward) practice phase (12 trials) to get familiar with the task, participants performed the main experiment in which

---

[1] While the term multimodal cuing *benefit* is used in the multimodal cuing literature, it is not optimal in the present task context, as it also entails performance impairment in invalidly cued trials.

[2] The term stimulus onset asynchronies (SOAs) always refers to the interval between cue onset and target onset (cue-target interval) in the present study. Cues and targets were separated by 100 ms at a minimum in accordance with typical Posner paradigms (for simultaneous cue-target presentation, see, e.g., Frassinetti et al., 2002; Lu et al., 2009). Multimodal cues were always presented simultaneously. This is different from studies exploring the boundary conditions of multisensory integration, which often feature variable SOAs between the multimodal signals themselves (see Diederich & Colonius, 2015).
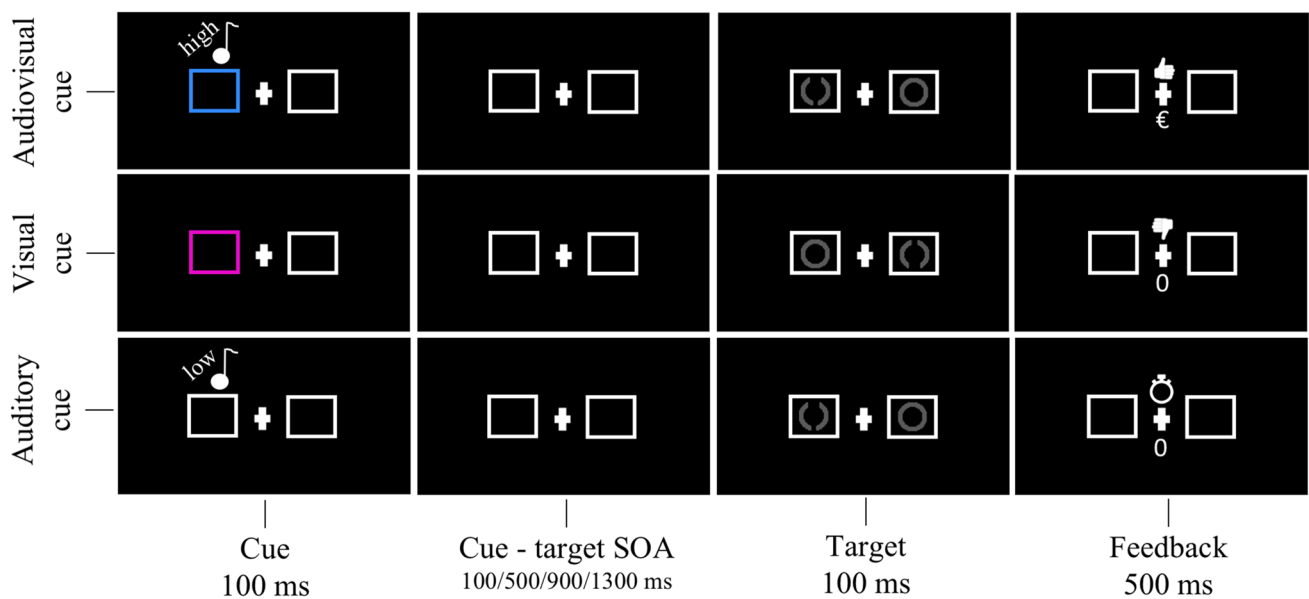
Fig. 1 Schematic representation of the trials in Experiment 1. At the start of the trials, an audiovisual, visual, or auditory cue predicted the location (left/right on the screen) of the upcoming target with a validity of about 80%. In this example, cues of blue color/high tone predicted reward prospect, while cues of pink color/low tone were not associated with reward. Subsequent targets consisted of two circles of which one has two unequally sized gaps. The participants' task was to identify the bigger of the two gaps (up/down) by means of button presses. Feedback was provided directly upon the response, signaling whether the response was correct and in-time (i.e., thumb-up icon), incorrect (i.e., thumb-down), or too late (i.e., clock). Of note, the words "high"/"low" and the musical notes serve for illustration only

they could earn additional bonus money (see below). During the presentation of all trial elements, a white fixation cross and two white place holders were visible on the screen. Each trial started with a cue presented for 100 ms that was either visual (blue/pink placeholder in left or right visual field; RGB = 50, 138, 255/RGB = 230, 10, 200), auditory (high/low tone in left or right ear via headphones; 2,500 Hz/1,000 Hz; 78 db; Boersma, 2001), or audiovisual (simultaneous presentation of a visual and auditory cue). Visual and auditory cuing procedures were conceptually similar to those applied by Bucker and Theeuwes (2016) and Schürmann et al. (2003). The location (left and right visual field/ear) of these cues predicted the spatial location of the upcoming target in the majority of trials. This was made explicit to participants prior to the practice phase. In addition, one of the two colors and one of the two tones signaled reward prospect, while the respective other color and tone were not associated with rewards. These reward-color and reward-tone mappings were counterbalanced across participants, and explicitly instructed after the practice phase. At this time, participants were also informed that they could win a maximal bonus of 5 € in the experiment and that the exact amount would depend on their performance. Moreover, it was made explicit that rewards would be exclusively earned in trials featuring reward cues (never after no-reward cues), and that the reward value per trial would be identical for the different cue types (visual, auditory, or audiovisual cues).

Finally, simultaneously presented visual and auditory signals (multimodal) were never conflicting with regard to both the spatial location and the reward prospect.

After a variable interval (cue-target SOA: 100/500/900/1,300 ms), two circles were presented in the left and right visual field for 100 ms. One circle contained two differently sized gaps (target), while the other circle was intact (distractor). The target was presented in the cued visual field in about 80% of the trials, i.e., presentation of the colored placeholder and/or tone were spatially congruent with the target. In the remaining ~20% of trials, the target was presented in the uncued visual field. Note that cue validity was relatively high (80% as compared to the prototypical 50% for peripheral cues) to ensure that participants stay engaged in the task – as a reward cue that only predicts location at chance might not be considered motivationally relevant. Moreover, in the multimodal cuing literature it has been argued that chance-level validity is not compatible with real-life situations (Ho et al., 2009). The participants' task was to quickly indicate which of the two gaps of the target circle (up/down) was largest by pressing one of two buttons on a keyboard (up/down arrow). After the response or once the maximum response window had elapsed, participants were presented with a feedback screen for 500 ms. The next trial was presented after a variable inter-trial interval between 500 and 1,000 ms. Feedback for correct and in-time responses consisted of a thumb-up icon just above

the fixation cross, while incorrect and too-late responses were followed by a thumb-down and clock, respectively. A simultaneously-presented euro sign or zero below fixation indicated whether the performance in this particular trial had led to a monetary bonus or not. The feedback screen did not contain information on the actual reward value since the bonus in a single trial can be perceived as fairly low and might reduce overall motivation (here, 5 €/576 reward trials ≈ 0.87 cents). Instead, participants received information on the accumulated reward value after each block (see below). The feedback in a given trial was dependent on response accuracy but also response time (RT) in that the response window was dynamically adjusted on the basis of participants' individual performance (Cornsweet, 1962). This procedure should motivate all participants to perform the task in a similar way and ameliorate strategic slowing (Heitz, 2014). Specifically, each condition's specific response window was calculated as the average RT in the first set of trials of that specific condition, and continuously adjusted if the response accuracy for that condition differed from 80% (cf. Wittmann et al., 2005). The maximum duration of each trial's time-out was restricted by the maximum period of the pre-set response window (i.e., 1,000 ms; for similar procedures, see Hoofs et al., 2019). Importantly, independent of this feedback procedure, all responses in a pre-set response window between 150 and 1,000 ms after target onset were considered for the analyses (i.e., responses exceeding the dynamically adjusted time-out of a given condition and hence followed by too-late feedback were still analyzed if RT was within the 1,000-ms range). After each block of 288 trials, participants received an overview showing the percentage of correct and in-time responses, and the amount of money earned so far.

Each participant performed a total of 1,152 trials, which were distributed over the conditions formed by combinations of four experimental factors: Cue type (audiovisual/visual/auditory), Validity (valid/invalid), Reward (reward/no-reward), and SOA (100/500/900/1,300 ms). The valid conditions contained 38 trials each and invalid ones consisted of ten trials. The trials were presented in four blocks of 288 trials that were separated by short breaks. Participants were instructed to keep their eyes on the fixation cross throughout the experiment. Monitoring during the practice phase and checks at different moments during the main experiment (with a webcam) showed that participants had no difficulties in maintaining fixation.

Data analysis Analyses were performed in JASP version 0.9.1.0 (JASP Team, 2018). Premature responses (i.e., < 150 ms) and responses slower than 1,000 ms (irrespective

of the adaptive time-out) were excluded from each dataset.[3] RT (correct trials only) and error-rate data were submitted to repeated-measures analyses of variance (rANOVAs), with within-subject factors Cue type (audiovisual/visual/auditory), Validity (valid/invalid), Reward (reward/no-reward), and SOA (100/500/900/1,300 ms). Tests with a p-value between .05 and .1 are consistently reported as (non-significant) trends, and tests with a p-value above .1 are considered non-significant. Trend-level F-tests are reported for completion, but not followed up by post hoc tests. Greenhouse-Geiser corrections were applied to the degrees of freedom and p-values in case sphericity assumptions were violated (i.e., significant Mauchly's test of sphericity), and the p-values of post hoc tests were corrected for the number of possible comparisons by using Bonferroni corrections (for more information, see Online Supplementary Material (OSM)). Please note that the effect size Cohen's *d* reported for the post hoc contrasts does not correct for multiple comparisons.

Results Experiment 1

RT data The RT data are depicted in Fig. 2. An overview of the RT condition means is provided in Table S1 of the OSM. The statistical output (including significant main effects and interactions as well as post hoc tests) is provided in Table S3 (OSM). As expected, responses were faster to validly cued targets as compared to invalidly cued ones (Validity: $F(2, 82) = 57.83$, $p < .001$; $\eta^2_p = .585$), and largely faster at longer SOAs (SOA: $F(2.19, 89.65) = 94.53$, $p < .001$; $\eta^2_p = .697$). We also observed a main effect of Cue type ($F(2, 82) = 24.88$, $p < .001$; $\eta^2_p = .378$), with faster responses after audiovisual cues ($t(41) = -5.38$, $p < .001$; $d = -0.830$) and auditory cues ($t(41) = -6.16$, $p < .001$; $d = -0.950$) as compared to visual cues, but no significant differences between audiovisual and auditory cues ($p > .2$). Importantly, these main effects were accompanied by several interactions. First, the validity effect (i.e., invalid minus valid) was amplified by reward as compared to no-reward trials, as indexed by an interaction between Reward and Validity ($F(1, 41) = 29.10$, $p < .001$; $\eta^2_p = .415$). Second, a Cue type by Validity interaction ($F(2, 82) = 6.82$, $p = .002$; $\eta^2_p = .143$) resulted from a significantly larger validity effect for audiovisual as compared to auditory cues ($t(41) = 3.57$, $p = .003$; $d = 0.551$), with no difference between visual and auditory or audiovisual cues (all $p > .1$). Third, there were two interactions involving SOA (Cue type × SOA: $F(6, 246) = 2.65$, $p = .017$; $\eta^2_p = .061$; Cue type × Reward × SOA: $F(6, 246)$

---

[3] As the percentage of trials with too-late responses is fairly low for each of the experiments (Experiment 1: 0.9%, Experiment 2: 1.5%; Experiment 3: 1.5%; calculated for the analyses-included participants), we consider the time-out of 1,000 ms as sufficiently large.
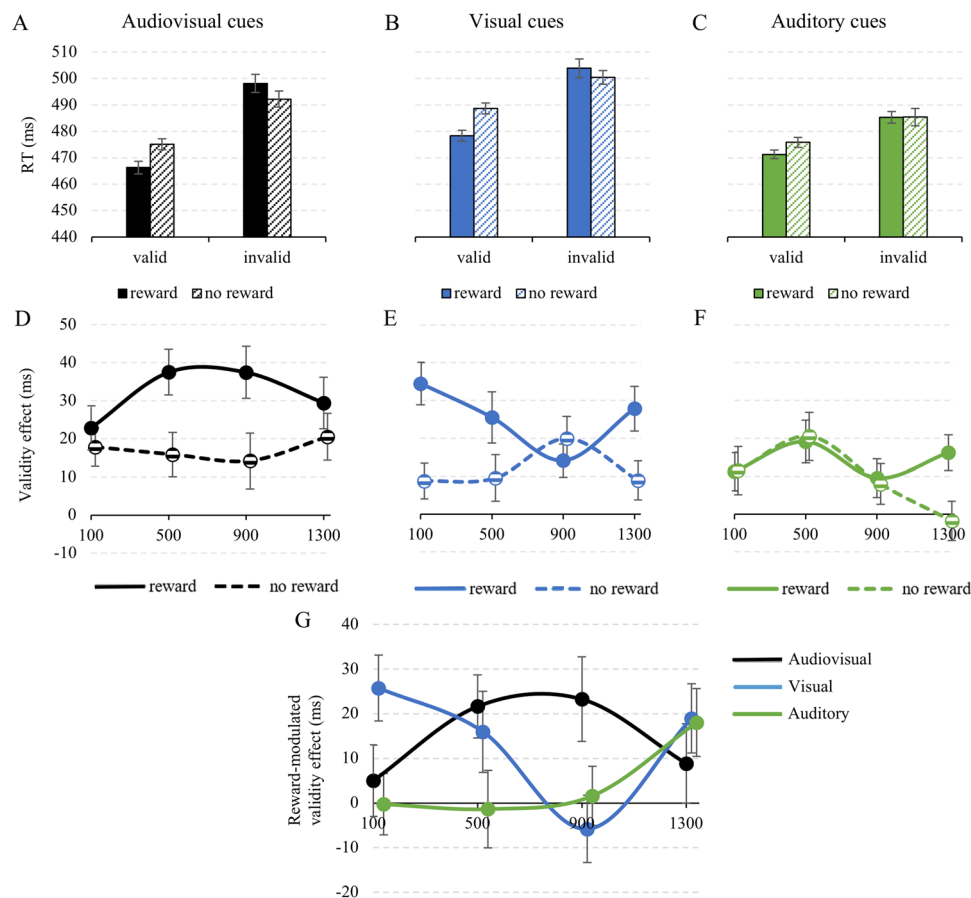
Fig. 2 Experiment 1: Response time (RT) data for audiovisual (**a** and **d**), visual (**b** and **e**), and auditory cue conditions (**c** and **f**). Solid bars/lines represent reward trials, while dashed bars/lines represent no-reward trials. In the middle graphs, the validity effect (i.e., invalid minus valid trials) is displayed for each individual stimulus-onset asynchrony (SOA) (100/500/900/1,300 ms). Panel G presents the reward-modulated validity effect (i.e., reward minus no-reward validity effect) and assists interpretation of the four-way interaction encompassing all factors. The subtle shifts in data points along the x-axes are merely illustrative (to avoid overlapping error bars). Error bars indicate ± 1 within-subject standard error (Cousineau-Morey method, for more information see O'Brien & Cousineau, 2015)

= 2.98, *p* = .008; η^2_p = .068). For the former, post hoc tests revealed that responses after visual cues were significantly slower compared to auditory cues at all SOAs (all *p* < .02), and significantly slower compared to audiovisual cues at SOAs of 100, 900, and 1,300 ms (all *p* < .007). The remaining post hoc tests were non-significant (all *p* > .06). For the three-way interaction, only one targeted post hoc test was significant, reflecting a larger reward-based response facilitation for visual compared to audiovisual cues at 900 ms (*p* = .008). The remaining post hoc tests within this interaction were non-significant (all *p* > .1).

Importantly, the above main effects and interactions were accompanied by a significant four-way interaction encompassing *all* experimental factors (Cue type × Reward × Validity × SOA: *F*(6, 246) = 2.65, *p* = .017; η^2_p = .061; Fig. 2d–f). Note that the effect size is well above the critical

effect size[4] of η^2_p = .050. In order to capture this complex interaction in a comprehensive way and in line with our research question, we created average double-difference scores representing the reward-modulated validity effect (Fig. 2g), and compared those statistically between Cue types within each SOA. These post hoc tests were corrected based on the number of possible tests within each SOA. We found that the reward-modulated validity effect was significantly larger for visual cues as compared to auditory ones at the shortest SOA (100 ms; *t*(41) = 2.60, *p* = .039; *d* = 0.401), with audiovisual cues in between (both *p* > .2). This pattern changed with increasing SOA, featuring a significantly larger reward-modulated validity effect for

---

[4] Critical effect sizes reflect the minimal effect size required for statistical significance. These are calculated as proposed by Lakens (2013).
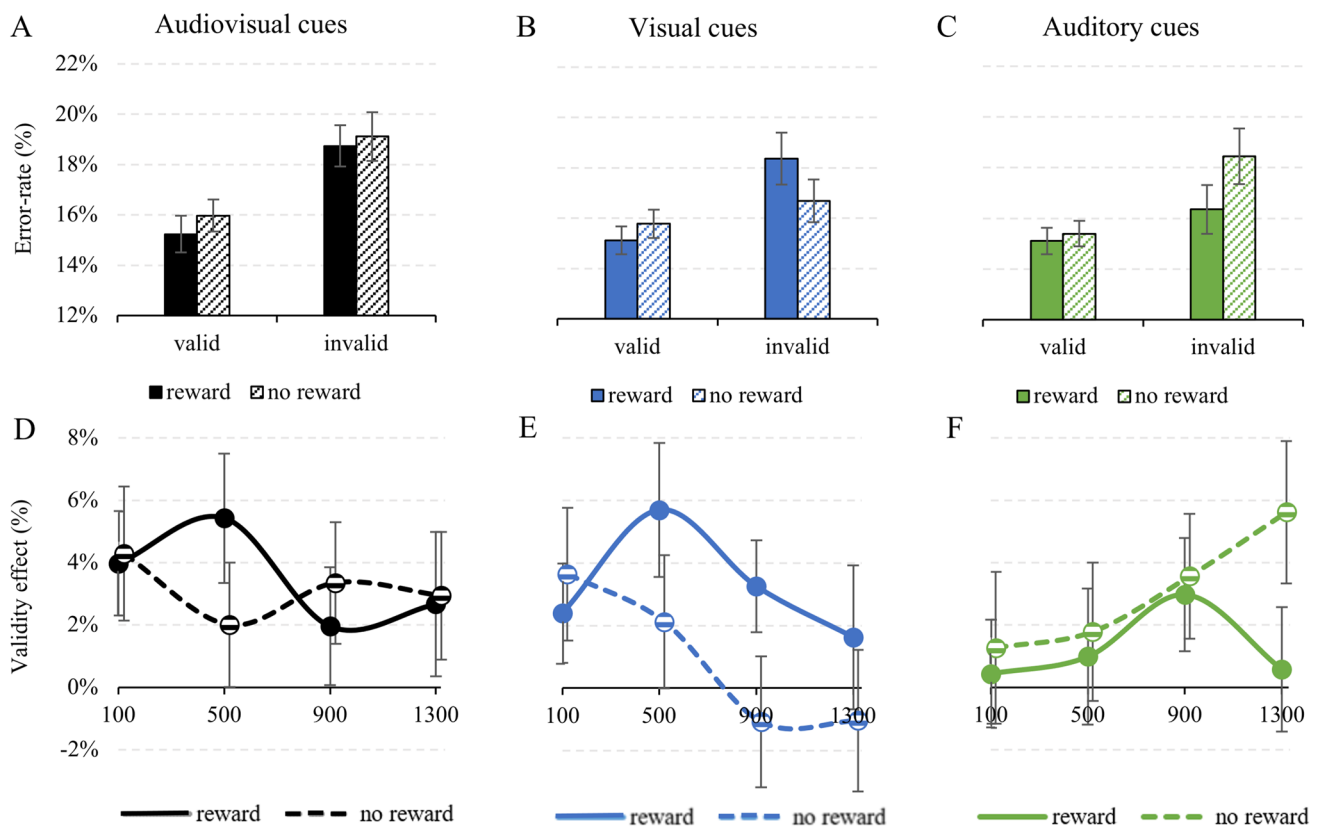
Fig. 3 Experiment 1: Error-rate data for audiovisual (**a** and **d**), visual (**b** and **e**), and auditory cue conditions (**c** and **f**). Solid bars/lines represent reward trials, while dashed bars/lines represent no-reward trials. In the bottom graphs, the validity effect (i.e., invalid minus valid trials) is displayed for each individual stimulus-onset asynchrony (SOA) (100/500/900/1,300 ms). As the four-way interaction encompassing all tested factors was non-significant in the error-rate data, there is no panel representing the reward-modulated validity effect (cf. response time (RT) data). The subtle shifts in data points along the x-axes are merely illustrative (to avoid overlapping error bars). Error bars indicate ± 1 within-subject standard error

audiovisual cues as compared to visual ones at an SOA of 900 ms ($t(41) = 2.72$, $p = .029$; $d = 0.419$), with auditory ones in between (both $p > .2$). There were no further significant post hoc tests in this targeted analysis, i.e., within the four-way interaction (all $p > .1$). The remaining main effects and interactions of the RT ANOVA were non-significant (all $p > .1$).

Error-rate data The error-rate data are depicted in Fig. 3. The error-rate condition means and the statistical output are provided in Tables S2 and S4 (OSM). Participants made more errors in trials featuring invalid compared to valid cues (Validity: $F(1, 41) = 12.36$, $p = .001$; $\eta^2_p = .232$). Furthermore, there was a trend-level interaction between Cue type, Reward, and Validity ($F(2, 82) = 3.04$, $p = .053$; $\eta^2_p = .069$), which is reported for completion but not further explored. The remaining main effects and interactions (including the four-way interaction with all factors) were non-significant (all $p > .1$). Of note, the effect size of the interaction with all factors ($p > .7$) yielded an effect size ($\eta^2_p = .014$) that was far below the critical value of .050.

Interim summary Experiment 1

The main results of Experiment 1 are that all cues lead to robust validity effects (invalid minus valid) in both RT and accuracy data, and that reward increased the validity effect in the RT data (Reward × Validity interaction). Moreover, visual and audiovisual reward cues differentially affected the validity effect in RT at different time points (as indexed by a four-way interaction with all factors). While the validity effect of visual reward cues was differentially increased at an SOA of 100 ms, the effect of multimodal reward cues was significantly pronounced at an SOA of 900 ms. Auditory reward cues did not feature a robust differential amplification of the validity effect at any of the SOAs. This pattern indicates that multimodal reward cues affect attentional guidance in a qualitatively different manner than unimodal (visual) reward cues, and is generally in line with a multimodal cuing benefit.[1] Based on the observation that visual and auditory reward signals seemed to affect attentional orienting differently in the present paradigm, we conducted two follow-up experiments featuring visual and auditory cues

separately. Specifically, we wanted to test whether these differences were due to superior visuospatial specificity of visual (reward) cues, in line with previous cross-modal cuing studies (see, e.g., Schürmann et al., 2003), or whether they reflect a context effect in that visual information might be more salient and overshadow auditory signals when presented in the same experiment. The reduced design allowed increasing the number of SOA steps to explore attentional orienting based on these different reward cues with higher temporal resolution.

Experiment 2

Methods Experiment 2

Participants The same recruitment procedure and inclusion criteria were applied as in Experiment 1. Data were collected from 50 students (45 females). The data of two participants were excluded due to a high number of errors (i.e., more than two standard deviations from the group mean). The remaining 48 participants (43 females) were on average 19.1 years old (SD = 2.8 years). In addition to one course credit, participants received an average bonus of 3.72 € (about 80% of 5 €).

Paradigm and procedure This experiment entailed only the visual cue conditions of Experiment 1, but with an extended set of SOAs between cue and target (100–1,500 ms) to allow for a more fine-grained description of the pattern, as well as a longer inter-trial interval (1,000–2,000 ms). The layout of the visual cues is described in Experiment 1. In order to create similar overall reward expectancy (albeit lower trial numbers), the bonus per reward trial was increased.

The total number of 768 trials per participant was distributed over conditions formed by combinations of the factors Validity (valid/invalid), Reward (reward/no-reward), and SOA (100/300/500/700/900/1,100/1,300/1,500). To match the 80:20% validity ratio of Experiment 1 with a different trial distribution, the valid conditions now contained 38 or 39 trials, and the invalid conditions nine or ten trials each. The amounts of valid and invalid trials within each participant were kept consistent, and the particular conditions containing the higher or lower number of trials per cell were counterbalanced across participants. The other settings and procedures were equivalent to Experiment 1.

Data analysis The same analysis procedures were applied as in Experiment 1. RT and error-rate data were submitted to 2 × 2 × 8 rANOVA, with within-subject factors Validity (valid/invalid), Reward (reward/no-reward), and SOA (100/300/500/700/900/1,100/1,300/1,500).

Results Experiment 2

RT data The RT data are depicted in Fig. 4 (panels a and b). An overview of the statistical output is provided in Table S5 (OSM). As in Experiment 1, responses were faster if the target location was validly cued (Validity: $F(1, 47) = 51.57$, $p < .001$; $\eta^2_p = .523$), and faster at longer SOAs (SOA: $F(5.16, 242.73) = 42.57$, $p < .001$; $\eta^2_p = .475$). The validity effect was more pronounced for reward as compared to no-reward trials as indexed by a significant interaction between Reward and Validity ($F(1, 47) = 8.48$, $p = .005$; $\eta^2_p = .153$; Fig. 4a), attesting that the difference between invalidly and validly cued targets was amplified in reward ($t(47) = 7.83$, $p < .001$; $d = 1.129$) compared to no-reward trials ($t(47) = 4.48$, $p < .001$; $d = 0.646$). There was no higher-order interaction with SOA in this experiment (Reward × Validity × SOA: $p > .1$; the η^2_p of .037 is below the critical effect size of .042). The remaining main effects and interactions were non-significant (all $p > .1$).

Error-rate data The error-rate data are depicted in Fig. 4 (panels c and d). An overview of the statistical output is provided in Table S6. In line with Experiment 1, we observed increased error-rates for invalidly compared to validly cued targets (Validity: $F(1, 47) = 11.39$, $p = .001$; $\eta^2_p = .195$). The validity effect in error-rate was not significantly modulated by Reward (Reward × Validity: $p > .5$). The remaining main effects and interactions were non-significant (all $p > .1$). Of note, the effect size of the three-way interaction encompassing all experimental factors ($p > .9$), $\eta^2_p = .008$, is substantially lower than the critical effect size of .042.

Interim summary Experiment 2

In a context with visual cues only, we found robust validity effects (in RT and accuracy) that were further enhanced in the RT data when the cue signaled reward compared to no-reward. This is in accordance with previous studies using peripheral visual reward cues (Bucker & Theeuwes, 2016), and resonates with the effects of visual cues in Experiment 1. In contrast to Experiment 1, the reward-modulated validity effect was not significantly pronounced at short versus long SOAs.

Experiment 3

Methods Experiment 3

Participants The same recruitment procedure and inclusion criteria were applied as in Experiments 1 and 2. Data were collected from 50 students (41 females). The data of four participants were excluded due to a high number of errors
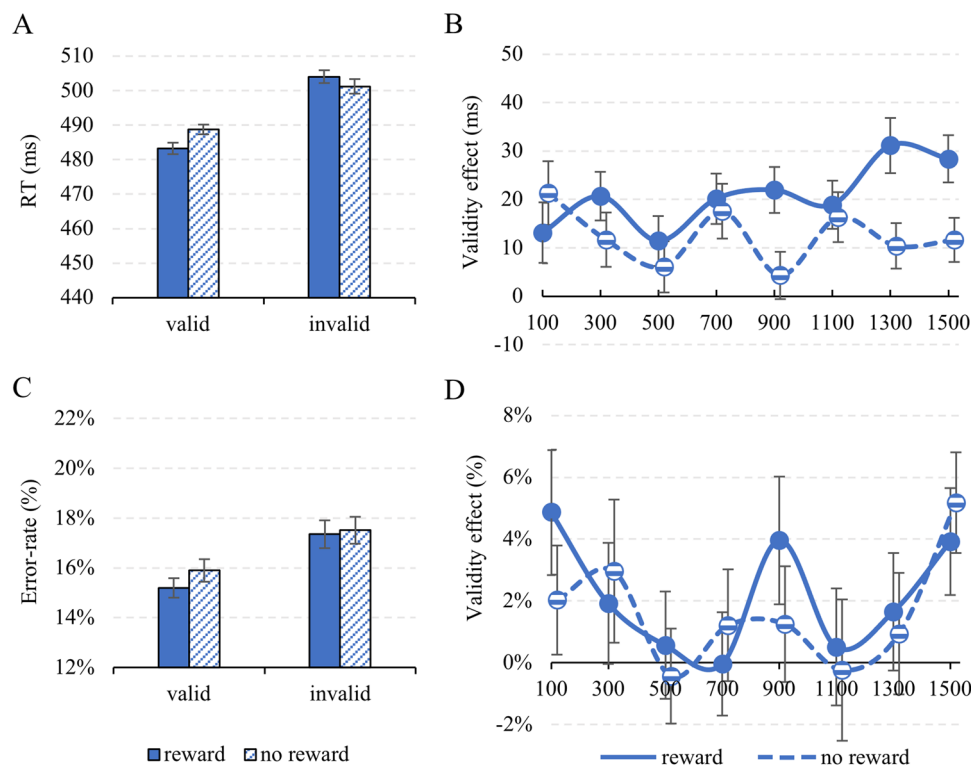
Fig. 4 Experiment 2: Response time (RT) (**a** and **b**) and error-rate (**c** and **d**) data for different cue conditions. Solid bars/lines represent reward trials, while dashed bars/lines represent no-reward trials. In the right graphs, the validity effect (i.e., invalid minus valid trials) is displayed for each individual stimulus-onset asynchrony (SOA) (10 0/300/500/700/900/1,100/1,300/1,500 ms). The subtle shifts in data points along the x-axes are merely illustrative (to avoid overlapping error bars). Error bars indicate ± 1 within-subject standard error

(i.e., more than two standard deviations from the group mean). The remaining 46 participants (37 females) were on average 18.4 years old (SD = 0.9 years). In addition to one course credit, participants received an average bonus of 3.74 € (about 80% of 5 €).

Paradigm and procedure This experiment was equivalent to Experiment 2 in all regards except that it featured auditory cues instead of visual ones. The parameters of the auditory cues are described in Experiment 1.

Data analysis Identical analysis procedures were applied as in Experiment 2.

Results Experiment 3

RT data The RT data are depicted in Fig. 5 (panels a and b). An overview of the statistical output is provided in Table S7 (OSM). As in Experiments 1 and 2, responses were faster for validly compared to invalidly cued targets (Validity: $F(1, 45)$ = 25.29, $p < .001$; $\eta^2_p = .360$), indicating that auditory cues are effective in guiding visual attention. Moreover, responses were faster at longer SOAs (SOA: $F(4.58, 206.13) = 20.13$, $p < .001$; $\eta^2_p = .309$). In contrast to Experiment 2 (featuring

visual instead of auditory cues), the validity effect was not significantly modulated by Reward (Reward × Validity: $p > .5$). Moreover, there was a trend-level main effect of Reward ($F(1, 45) = 3.17$, $p = .082$; $\eta^2_p = .066$), as well as a trend-level interaction between Validity and SOA ($F(7, 315) = 1.78$, $p = .090$; $\eta^2_p = .038$). These trends are reported for completion but not further explored. There was no significant higher-order interaction in this experiment (Reward × Validity × SOA: $p > .9$; the η^2_p of .006 is clearly lower than the critical effect size of .043). The remaining main effects and interactions were non-significant (all $p > .2$).

Error-rate data The error-rate data are depicted in Fig. 5 (panels c and d). An overview of the statistical output is provided in Table S8 (OSM). Again, participants committed more errors when responding to invalidly compared to validly cued targets (Validity: $F(1, 45) = 6.95$, $p = .011$; $\eta^2_p = .134$), once more confirming that auditory cues can guide visual attention. In line with the RT data of this experiment, the validity effect in error-rate was not significantly modulated by Reward (Reward × Validity: $p > .7$). The remaining main effects and interactions were non-significant (all $p > .1$). As in the RT data, the interaction between Reward,
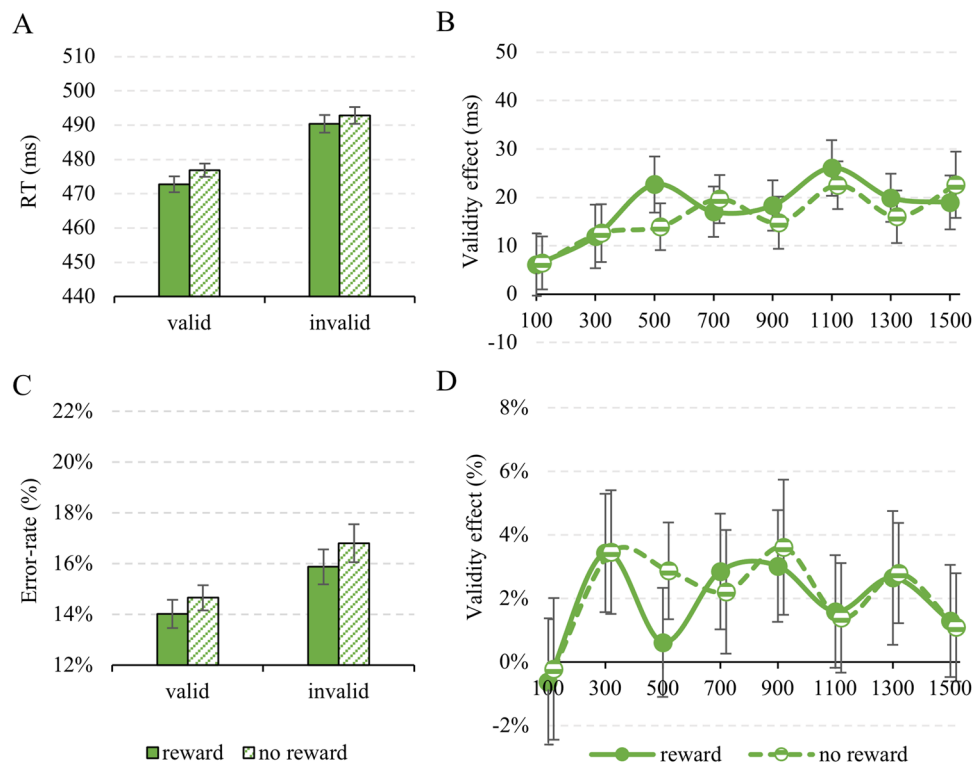
Fig. 5 Experiment 3: Response times (RTs) (**a** and **b**) and error-rate (**c** and **d**) data for different cue conditions. Solid bars/lines represent reward trials, while dashed bars/lines represent no-reward trials. In the right graphs, the validity effect (i.e., invalid minus valid trials) is displayed for each individual stimulus-onset asynchrony (SOA) (10 0/300/500/700/900/1,100/1,300/1,500 ms). The subtle shifts in data points along the x-axes are merely illustrative (to avoid overlapping error bars). Error bars indicate ± 1 within-subject standard error

Validity, and SOA was non-significant ($p > .9$; the η^2_p of .003 is clearly below the critical effect size of .043).

Interim summary Experiment 3

In a task with auditory cues alone, we found robust validity effects (RT and accuracy), which were, however, not further modulated by reward (in contrast to Experiment 2). Moreover, contrary to Experiment 1, which featured a significantly pronounced validity effect at 500 ms for auditory cues, this interaction (Validity × SOA) was only trending here, with a numerical increase of the validity effect with longer SOAs. For completion, the main effect of reward was also only trending in this experiment. Together, in a context with auditory cues only, the prospect of reward did not have a significant effect on performance, neither generally nor on the validity effect in particular.

General discussion

Results summary

Inspired by the rich literature on reward-based enhancement of attention on the one hand and multimodal cuing studies

on the other, we aimed to investigate how simultaneously presented visual and auditory reward signals would guide attention. To this end, we employed unimodal visual and auditory as well as multimodal audiovisual cues in a Posner task (Experiment 1). In addition to indicating the spatial location of the upcoming target (with a probability of 80%), these cues signaled the prospect of reward on a given trial. Overall, across cue types (audiovisual, visual, and auditory) and SOAs, we found robust validity effects (invalid minus valid) in both RT (Fig. 2a–c) and accuracy (Fig. 3a–c) measures, as well as an amplification of the validity effect by reward prospect in RT. While this reward-modulated validity effect seemed numerically larger for visual and audiovisual cues as compared to auditory ones (Fig. 2d–f), there was no statistical support for this notion. Intriguingly, however, we found differences between the impact of visual, auditory, and audiovisual reward cues on attentional guidance depending on the length of the cue-target interval (i.e., SOA), as signified by a four-way interaction between all factors (Fig. 2g). While the effect of visual reward cues on response speed was differentially larger at short SOAs (in particular at 100 ms), the effect of audiovisual cues emerged at longer SOAs (in particular at 900 ms). Notably, auditory reward cues did not differentially increase the validity effect at any SOA. To

further explore the differential effects of visual and auditory reward cues, and hence their contribution to the multimodal effect in the main experiment, we conducted two follow-up experiments featuring unimodal visual and auditory (reward) cues in isolation (Experiments 2 and 3). These experiments revealed that both visual and auditory cues were effective in guiding visual attention, as indexed by validity effects in RT and accuracy data, but only the visual experiment featured an additional modulation of the RT validity effect by reward. In contrast to Experiment 1, the reward-modulated validity effect was not significantly influenced by SOA in the follow-up experiments. Finally, there was a tendency for faster responses in reward trials across conditions in the auditory experiment; however, the effect was only trending. In what follows, we discuss the observations of the multimodal experiment, as well as the unimodal results, followed by a general conclusion section.

The effect of multimodal reward cues (Experiment 1)

To investigate the effect of simultaneously presented visual and auditory reward cues on visual attention, we combined elements from two separate research lines, i.e., one studying reward effects on visual attention (e.g., Bucker & Theeuwes, 2016; Engelmann & Pessoa, 2007; Krebs et al., 2012; Small et al., 2005) and one focusing on attentional guidance by multimodal signals (e.g., Barrett & Krumbholz, 2012; Frassinetti et al., 2002; Spence & Driver, 1999). Globally, our observations are well in line with previous literature in both research areas. First, previous research has shown that cues signaling the prospect of reward affect performance in visual attention tasks, including improved visual discrimination at validly cued target locations (e.g., Bucker & Theeuwes, 2014, 2016; Engelmann & Pessoa, 2007; Krebs et al., 2012; Munneke et al., 2015; Schevernels et al., 2014; Small et al., 2005), but also performance costs at invalidly cued locations based on enhanced attentional capture (e.g., Bucker & Theeuwes, 2014, 2016; Munneke et al., 2015). This latter observation is also in line with more incidental (i.e., non-strategic) effects of reward-related stimulus features, which can capture attention if they are presented as distractors in a reward task (e.g., Hickey et al., 2010; Itthipuripat et al., 2019), a non-rewarded test phase (e.g., Anderson, 2013; Anderson et al., 2011; Watson et al., 2019), or in both the reward and non-rewarded phase of the same experiment (see Watson et al., 2020).

Second, in the multimodal cuing literature it has been demonstrated that presenting two simultaneous signals from different modalities can enhance attentional orienting (McDonald et al., 2000; McDonald & Ward, 2000; Santangelo, Ho, & Spence, 2008a). While multimodal reward cues did not increase the validity effect in magnitude, they featured a different temporal profile as compared to unimodal reward cues (Fig. 2g). Specifically, the differential validity effect of audiovisual reward cues emerged later in the cue-target interval (and hence seemed more extended), which was in contrast to an early validity effect of visual reward cues. Considering the predictions regarding the joint effect of multiple reward signals on attentional guidance, our results are neither entirely in line with an additive effect (linear integration hypothesis), nor entirely in line with the saliency notion (maximum hypothesis), but are indicative of a qualitative change of attentional guidance. We further discuss the nature and implications of the observed effects in the *Conclusions* section, after considering the results of the unimodal experiments.

Two additional observations seem relevant for the interpretation of the data. First, visual cues slowed down responses as compared to both auditory and audiovisual cues, which may be the result of impaired disengagement from the cued location (Watson et al., 2020) and/or stronger inhibition of return (Bucker & Theeuwes, 2014). Interestingly, this was less pronounced if an additional auditory signal was present (multimodal condition). Second, judging from the means of the validity effect over time (Fig. 2d–f), it seems that the reward-related validity effect increases again at 1,300 ms in both the visual and auditory condition. This effect (which is part of the four-way interaction) resonates with a more strategic (or endogenous) form of attentional orienting at longer SOAs (Godijn & Theeuwes, 2002). Such strategic orienting might be especially pronounced in the present paradigm because rewards are at stake and because cue predictability is relatively high (80%).

The effect of unimodal visual and auditory reward cues (Experiments 2 and 3)

In order to better understand the differential effects of visual and auditory reward signals in Experiment 1, and in particular with regard to their contribution to the multimodal effect, we now discuss the unimodal experiments (Experiments 2 and 3). We found that both visual (Fig. 4) and auditory (Fig. 5) cues modulated attentional orienting as indexed by robust validity effects in RT and accuracy data. This replicates previous work using intra-modal (same-modality) and cross-modal (different-modality) cues (Newport & Howarth, 2009; Schmitt et al., 2000; Spence & Driver, 1997). However, other studies in the field have found an advantage of visual over auditory cues for discriminating visual targets (Schürmann et al., 2003; Ward, 1994), which has been attributed to a benefit at the sensory processing level. Interestingly, one study found that attentional guidance is comparable for visual and auditory cues when presented in close temporal proximity to the target (Störmer et al., 2019).

With regard to the impact of reward prospect, the results differed between Experiments 2 and 3. In the visual

experiment, we found that the validity effect was enhanced by reward, replicating previous work (see Bucker & Theeuwes, 2016; Munneke et al., 2015) employing visual reward cues in a Posner task (but see Baines et al., 2011, for a study that merely found independent reward and validity effects). In contrast to this, the auditory experiment did not feature robust effects of reward prospect, neither in terms of the validity effect, nor globally (the main effect of reward was only trending). This difference may be attributed to a putative perceptual processing benefit of visual cues (e.g., Ward, 1994), which is additionally modulated by reward (e.g., Anderson, 2013; Hickey et al., 2010; Itthipuripat et al., 2019). This is also in line with the notion of a spatial saliency map (Itti & Koch, 2000; Itti & Koch, 2001), here modulated by reward information, which can in turn improve target discrimination at the respective location. Together, we suggest that the differential effect of visual reward cues in our experiment arises through the combination of this assumed sensory processing benefit of visual over auditory cues and reward-induced changes in the spatial saliency map. In other words, the salient reward information increases the sensory processing benefit of visual cues – which might not surface in all the before-mentioned studies due to paradigmatic differences and/or statistical power.

Intriguingly, the visual experiment did not replicate the modulations based on SOA that we observed in Experiment 1 (i.e., pronounced reward-modulated validity effect at the shortest SOA). The absence of this interaction with SOA may be due to the distribution of SOAs in that a higher number of intervals reduce temporal predictability of the upcoming target and in turn modulate strategic effects in attentional orienting (Nobre & van Ede, 2018). This null finding could also reflect a failure of replication (Maxwell et al., 2015), potentially related to low trial numbers in the individual cells of the higher-order interactions. This being said, the fact that the interaction between Validity, Reward, and SOA in the visual condition of Experiment 1 replicated an earlier study that featured only two SOA steps (Bucker & Theeuwes, 2016), suggests that temporal predictability plays a critical role here. That said, there is at least a numerical indication that strategic attentional orienting at the longest SOAs is modulated by reward (Fig. 4b), similar to the pattern in Experiment 1, but the interaction between Reward, Validity, and SOA was not significant.

This is, to the best of our knowledge, the first direct comparison between visual and auditory reward cues in an attentional-orienting paradigm. An interesting follow-up route would be to match the settings between auditory and visual cues even more. For instance, by presenting auditory signals via loudspeakers to the left and right of the screen (instead of headphones), they might be perceived as more spatially overlapping with the subsequently presented visual targets (for such a cuing procedure see, e.g., Talsma et al.,

2009). Finally, it is also possible that participants can more easily associate reward information with a certain modality in the first place (here, visual). While the absence of a robust reward main effect in the unimodal auditory experiment seems to support this notion, more data are needed to confirm this idea in that reward is directly bound to a spatial location in the present paradigm, leading to complex interactions between spatial attentional orienting and reward processing.

Together, with respect to the unimodal cue conditions of Experiment 1, the two follow-up experiments confirmed that visual reward cues have a larger impact in guiding attentional orienting towards visual targets, and that this difference is not merely due to the multimodal cueing context in Experiment 1 (in the sense that visual reward cues would overshadow auditory ones only if they are presented in the same task).

Conclusions

Returning to the main research question, i.e., the effect of multimodal reward cues on visual attention, it seems that the combination of visual and auditory reward cues leads to a qualitatively different process of attentional orienting, which is not simply the sum of the respective unimodal cuing effects. While auditory reward cues did not facilitate attentional guidance in a robust manner, they seemed to modulate the effect of visual reward cues. Importantly, the multimodal reward effect is not (significantly) larger than the unimodal visual one, but features a different temporal profile. Specifically, it seems that the initial orienting triggered by visual reward cues, likely reflecting saliency-based (bottom-up) capture, is replaced by a later (or more sustained) effect when an additional auditory reward signal is presented (see Fig. 2g). This pattern could be globally interpreted as a multimodal cuing benefit[5] (e.g., Frassinetti et al., 2002; Noesselt et al., 2008; Santangelo, Ho, Spence, 2008). That said, the specific modulation (i.e., the shift/extension of the validity effect in time) is likely generated by the nature of the current paradigm. First, as evidenced by the unimodal cuing conditions and related literature, visual and auditory signals have a differential impact on attentional orienting in a visual task. The higher spatial specificity of visual (reward) cues leads to stronger attentional capture, which is most pronounced at short SOAs (Bucker & Theeuwes, 2016). Second, recent work has shown that reward information not only captures attention more readily, but hinders disengagement away from the cued location if presented in close temporal proximity

---

[5] While the term multimodal cuing *benefit* is used in the multimodal cuing literature, it is not optimal in the present task context, as it also entails performance impairment in invalidly cued trials.

to the target (Watson et al., 2020). Third, the more strategic (endogenous) orienting of attention that typically emerges only at longer SOAs (Born et al., 2011) is likely comparable between visual and auditory (reward) cues, as this does not rely on attentional capture at a specific spatial location. Considering these observations and the differential temporal profiles of the reward-modulated validity effect in Fig. 2g, it might be that the additional auditory signal reduces the initial bottom-up capture by visual reward cues (and also the lingering at the cued location), and instead amplifies the more strategic component of attentional orienting. A potential reason why this is not happening at short cue-target SOAs could be a processing bottleneck (Broadbent, 1958) when presenting two salient signals simultaneously.

In this context, it is important to consider that the current paradigm likely emphasizes strategic effects at longer SOAs as compared to a typical Posner paradigm with peripheral cues. Specifically, the cue validity of 80% as well as the prospect of reward in half of the trials renders the cues more relevant and might promote strategic orienting to the cued location even after re-centering of attention. All of this being said, while increased cue relevance will likely promote strategic processes at longer SOAs, this will be the case for all conditions, and, most importantly, this is unlikely to affect the initial orienting process. A recent electro-encephalography study directly compared the effects of peripheral (exogenous) auditory cues with 50% versus 80% validity, as well as central (endogenous) ones on attentional orienting (Keefe & Störmer, 2021). First, all task versions yielded comparable validity effects on the behavioral level, and second, neural signatures of attentional orienting were similar for both peripheral cuing versions (50% and 80%), but different from the central cuing version, in that they emerged earlier in time. Moreover, previous work using visual reward cues with 50% validity reported fairly similar results to the present study, including comparable condition means (Bucker & Theeuwes, 2016).

Finally, we emphasize that the differential effect of multimodal reward cues is only observed when considering the SOA factor (in that instead of the hypothesized three-way interaction between Reward, Validity, and Cue type, we found a four-way interaction with SOA). This is not to say that there is no difference in the reward-modulated validity effect between cue types, but that the cue-target interval modulates the relationship between the other factors. This finding, together with the discrepancy in SOA effects in the visual condition of Experiments 1 and 2, calls for future investigations into the role of the cue-target interval. For instance, it would be valuable to replicate the current study using only two SOAs (short vs. long), which might emphasize differences between conditions even more due to the increased temporal predictability of the target onset.

The present results are not only informative from an experimental research perspective, but are also relevant for real life. While daily-life situations in which motivationally relevant information is limited to one modality are rare, research on reward effects on attentional orienting almost exclusively feature unimodal cues. Hence, exploring the effects of multimodal reward signals is highly relevant for industry and policy makers. Previous research has provided evidence that multimodal warning signals are more effective than unimodal ones (Haas & Van Erp, 2014; Oskarsson et al., 2012), especially if the signals were inherently linked to the warning situation (Graham, 1999) and mildly negative (Bellettiere et al., 2014). However, as negative warnings can be perceived as being too aversive, they might be counterproductive in that they cause disengagement and frustration (Bellettiere et al., 2014). Therefore, the use of positive signals opens up a new and interesting avenue for creation of safety signals. And in addition, studying and using positive multimodal signals seems consistent with the general contemporary notion that positive motivation is the more promising approach to change individuals' behavior in the long term (Armellino et al., 2012, 2013). That said, the opposite side of the coin is that these signals can be used to facilitate maladaptive behavior as well. Examples include TV commercials that make use of audiovisual signals to promote unhealthy food and drinks, but also the set-up of gambling facilities and video games, all of which can promote addictive behaviors.

To conclude, we found that audiovisual reward cues changed the time course of attentional orienting towards visual targets, especially as compared to unimodal visual reward cues. Instead of displaying a larger amplitude, the validity effect was shifted/extended in time when adding an auditory reward cue to the inherently dominant visual one. While being in line with a multimodal cuing benefit, this specific pattern highlights that two unimodal reward signals are not simply combined in a linear fashion but lead to a qualitatively different process. The differential temporal profiles seem to suggest that the additional auditory signal might reduce the strong attentional capture of visual cues and instead emphasizes more strategic orienting to the cued location. This interpretation is tentative and should be refined in future research.

CRediT authorship contribution statement

VH: Methodology; software; formal analysis; investigation; data curation; writing – original draft; writing – review and editing; visualization; project administration. IG: Methodology; formal analysis; investigation; writing – review and editing. CNB: Writing – review and editing; RMK: Conceptualization; methodology; writing – original draft; writing – review and editing; supervision.

Data availability De-identified data and statistical output will be made available after publication via this Open Science Framework page: https://osf.io/q4tja/

Other study materials are available upon request via the corresponding author.

Declaration

Competing interests The authors have no competing interests to declare.

Ethics approval All experiments of the present study are in accordance with the general ethical protocol of the local ethics board. Written informed consent was obtained from all participants in advance of participation. All procedures were in accordance with the Declaration of Helsinki from 1964 and its later amendments.

References

Anderson, B. A. (2013). A value-driven mechanism of attentional selection. *Journal of Vision, 13*(3), 1-16. https://doi.org/10.1167/13.3.7

Anderson, B. A. (2016). The attention habit: How reward learning shapes attentional selection. *Annals of the New York Academy of Sciences*, *1369*(1), 24–39. https://doi.org/10.1111/nyas.12957

Anderson, B. A., Laurent, P. A., & Yantis, S. (2011). Value-driven attentional capture. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(25), 10367. https://doi.org/10.1073/PNAS.1104047108

Armellino, D., Hussain, E., Schilling, M. E., Senicola, W., Eichorn, A., Dlugacz, Y., & Farber, B. F. (2012). Using high-technology to enforce low-technology safety measures: The use of third-party remote video auditing and real-time feedback in healthcare. *Clinical Infectious Diseases, 54*(1), 1-7. https://doi.org/10.1093/cid/cir773

Armellino, D., Trivedi, M., Law, I., Singh, N., Schilling, M. E., Hussain, E., & Farber, B. (2013). Replicating changes in hand hygiene in a surgical intensive care unit with remote video auditing and feedback. *American Journal of Infection Control, 41*(10), 925-927. https://doi.org/10.1016/j.ajic.2012.12.011

Baines, S., Ruz, M., Rao, A., Denison, R., & Nobre, A. C. (2011). Modulation of neural activity by motivational and spatial biases. *Neuropsychologia*. https://doi.org/10.1016/j.neuropsychologia.2011.04.029

Barrett, D. J. K., & Krumbholz, K. (2012). Evidence for multisensory integration in the elicitation of prior entry by bimodal cues. *Experimental Brain Research, 222*(1-2), 11-20. https://doi.org/10.1007/s00221-012-3191-8

Bellettiere, J., Hughes, S. C., Liles, S., Boman-Davis, M., Klepeis, N., Blumberg, E., … Hovell, M. F. (2014). Developing and selecting auditory warnings for a real-time behavioral intervention. *American Journal of Public Health Research, 2*(6), 232-238. https://doi.org/10.12691/ajphr-2-6-3

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International*, *5*(9/10), 341-347. https://doi.org/10.1097/AUD.0b013e31821473f7

Born, S., Kerzel, D., & Theeuwes, J. (2011). Evidence for a dissociation between the control of oculomotor capture and disengagement. *Experimental Brain Research*. https://doi.org/10.1007/s00221-010-2510-1

Broadbent, D. E. (1958). *Perception and communication*. Pergamon Press.

Bucker, B., & Theeuwes, J. (2014). The effect of reward on orienting and reorienting in exogenous cuing. *Cognitive, Affective and Behavioral Neuroscience*, *14*(2), 635-646. https://doi.org/10.3758/s13415-014-0278-7

Bucker, B., & Theeuwes, J. (2016). Appetitive and aversive outcome associations modulate exogenous cueing. *Attention, Perception, and Psychophysics, 78*, 2253–2265. https://doi.org/10.3758/s13414-016-1107-6

Camara, E., Manohar, S., & Husain, M. (2013). Past rewards capture spatial attention and action choices. *Experimental Brain Research*, *230*(3), 291-300. https://doi.org/10.1007/s00221-013-3654-6

Chelazzi, L., Perlato, A., Santandrea, E., & Della Libera, C. (2013). Rewards teach visual selective attention. *Vision Research*, *85*, 58–62. https://doi.org/10.1016/j.visres.2012.12.005

Cornsweet, T. N. (1962). The staircase-method in psychophysics. *The American Journal of Psychology*, *75*(3), 485-491. https://doi.org/10.2307/1419876

Diederich, A., & Colonius, H. (2015). The time window of multisensory integration: Relating reaction times and judgments of temporal order. *Psychological Review*, *122*(2), 232–241. https://doi.org/10.1037/A0038696

Driver, J. (2001). A selective review of selective attention research from the past century. *British Journal of Psychology, 92*(1), 53-78. https://doi.org/10.1348/000712601162103

Driver, J., & Noesselt, T. (2008). Multisensory interplay reveals cross-modal influences on "sensory-specific" brain regions, neural responses, and judgments. *Neuron, 57*(1), 11-23. https://doi.org/10.1016/j.neuron.2007.12.013

Engelmann, J. B., & Pessoa, L. (2007). Motivation sharpens exogenous spatial attention. *Emotion, 7*(3), 668-674. https://doi.org/10.1037/1528-3542.7.3.668

Failing, M., & Theeuwes, J. (2018). Selection history: How reward modulates selectivity of visual attention. *Psychonomic Bulletin and Review, 25*(2), 514-538. https://doi.org/10.3758/s13423-017-1380-y

Frassinetti, F., Bolognini, N., & Làdavas, E. (2002). Enhancement of visual perception by crossmodal visuo-auditory interaction. *Experimental Brain Research*, *147*(3), 332-343. https://doi.org/10.1007/s00221-002-1262-y

Godijn, R., & Theeuwes, J. (2002). *Programming of Endogenous and Exogenous Saccades: Evidence for a Competitive Integration Model*. https://doi.org/10.1037/0096-1523.28.5.1039

Graham, R. (1999). Use of auditory icons as emergency warnings: Evaluation within a vehicle collision avoidance application. *Ergonomics, 42*(9), 1233-1248. https://doi.org/10.1080/001401399185108

Haas, E. C., & Van Erp, J. B. F. (2014). Multimodal warnings to enhance risk communication and safety. *Safety Science*, *61*, 29-35. https://doi.org/10.1016/j.ssci.2013.07.011

Heitz, R. P. (2014). The speed-accuracy tradeoff: History, physiology, methodology, and behavior. *Frontiers in Neuroscience, 8,* 150. https://doi.org/10.3389/fnins.2014.00150

Hickey, C., Chelazzi, L., & Theeuwes, J. (2010). Reward changes salience in human vision via the anterior cingulate. *Journal of

Neuroscience, *30*(33), 11096–11103. https://doi.org/10.1523/jneurosci.1026-10.2010

Ho, C., Santangelo, V., & Spence, C. (2009). Multisensory warning signals: when spatial correspondence matters. *Experimental Brain Research 2009 195:2*, *195*(2), 261–272. https://doi.org/10.1007/S00221-009-1778-5

Hoofs, V., Carsten, T., Boehler, C. N., & Krebs, R. M. (2019). Interactions between incentive valence and action information in a cued approach–avoidance task. *Psychological Research, 83*(1), 13-25. https://doi.org/10.1007/s00426-018-0975-x

Itthipuripat, S., Vo, V. A., Sprague, T. C., & Serences, J. T. (2019). Value-driven attentional capture enhances distractor representations in early visual cortex. *PLoS Biology, 17*(8). https://doi.org/10.1371/journal.pbio.3000186

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research, 40*(10-12), 1489-1506. https://doi.org/10.1016/S0042-6989(99)00163-7

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience, 2*(3), 194-203. https://doi.org/10.1038/35058500

JASP Team. (2018). JASP (Version 0.9)[Computer software]. *[Computer Software]*.

Keefe, J. M., & Störmer, V. S. (2021). Lateralized alpha activity and slow potential shifts over visual cortex track the time course of both endogenous and exogenous orienting of attention: Endogenous & exogenous orienting of attention. *NeuroImage*, *225*, 117495. https://doi.org/10.1016/j.neuroimage.2020.117495

Klemen, J., & Chambers, C. D. (2012). Current perspectives and methods in studying neural mechanisms of multisensory interactions. *Neuroscience and Biobehavioral Reviews*. https://doi.org/10.1016/j.neubiorev.2011.04.015

Krebs, R. M., Boehler, C. N., Roberts, K. C., Song, A. W., & Woldorff, M. G. (2012). The involvement of the dopaminergic midbrain and cortico-striatal-thalamic circuits in the integration of reward prospect and attentional task demands. *Cerebral Cortex*, *22*(3), 607-615. https://doi.org/10.1093/cercor/bhr134

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *0*(NOV), 863. https://doi.org/10.3389/FPSYG.2013.00863

Lu, Z. L., Tse, H. C. H., Dosher, B. A., Lesmes, L. A., Posner, C., & Chu, W. (2009). Intra- and cross-modal cuing of spatial attention: Time courses and mechanisms. *Vision Research, 49*(10), 1081-1096. https://doi.org/10.1016/j.visres.2008.05.021

Maxwell, S. E., Lau, M. Y., & Howard, G. S. (2015). Is psychology suffering from a replication crisis?: What does "failure to replicate" really mean? *American Psychologist*, *70*(6), 487-498. https://doi.org/10.1037/a0039400

McDonald, J. J., & Ward, L. M. (2000). Involuntary listening AIDS seeing: Evidence from Human Electrophysiology. *Psychological Science*, *11*(2), 167-171. https://doi.org/10.1111/1467-9280.00233

McDonald, J. J., Teder-Saälejärvi, W. A., & Hillyard, S. A. (2000). Involuntary orienting to sound improves visual perception. *Nature, 407*(6806). https://doi.org/10.1038/35038085

McDonald, J. J., Teder-Sälejärvi, W. A., & Ward, L. M. (2001). Multisensory integration and crossmodal attention effects in the human brain. *Science*, *292*(5523), 1791. https://doi.org/10.1126/SCIENCE.292.5523.1791A

Meredith, M. A. , & Stein, B. E. (1986). Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *Journal of Neurophysiology*, *56*(3), 640–662. https://doi.org/10.1152/JN.1986.56.3.640

Milstein, D. M., & Dorris, M. C. (2007). The influence of expected value on saccadic preparation. *Journal of Neuroscience*, *27*(18), 4810-4818. https://doi.org/10.1523/JNEUROSCI.0577-07.2007

Mobbs, D., Hagan, C. C., Dalgleish, T., Silston, B., & Prévost, C. (2015). The ecology of human fear: Survival optimization and the nervous system. *Frontiers in Neuroscience, 9*, 55. https://doi.org/10.3389/fnins.2015.00055

Munneke, J., Hoppenbrouwers, S. S., & Theeuwes, J. (2015). Reward can modulate attentional capture, independent of top-down set. *Attention, Perception, and Psychophysics*, *77*(8), 2540-2548. https://doi.org/10.3758/s13414-015-0958-6

Newport, R., & Howarth, S. (2009). Social gaze cueing to auditory locations. *Quarterly Journal of Experimental Psychology, 62*(4), 625-634. https://doi.org/10.1080/17470210802486027

Ngo M. K, & Spence, C. (2010). Auditory, tactile, and multisensory cues facilitate search for dynamic visual stimuli. *Attention, Perception & Psychophysics*, *72*(6), 1654–1665. https://doi.org/10.3758/APP.72.6.1654

Nobre, A. C., & van Ede, F. (2018). Anticipated moments: temporal structure in attention. *Nature Reviews Neuroscience,* 19(1), 34-48. https://doi.org/10.1038/nrn.2017.141

Noesselt, T., Bergmann, D., Hake, M., Heinze, H. J., & Fendrich, R. (2008). Sound increases the saliency of visual events. *Brain Research, 1220,* 157-163. https://doi.org/10.1016/j.brainres.2007.12.060

O'Brien, F., & Cousineau, D. (2015). Representing error bars in within-subject designs in typical software packages. *The quantitative methods for psychology*, *10*(1), 56-67.

Oskarsson, P. A., Eriksson, L., & Carlander, O. (2012). Enhanced perception and performance by multimodal threat cueing in simulated combat vehicle. *Human Factors, 54*(1), 122-137. https://doi.org/10.1177/0018720811424895

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*(1), 3-25. https://doi.org/10.1080/00335558008248231

Quigley, C., Onat, S., Harding, S., Cooke, M., & König, P. (2007). Audio-visual integration during overt visual attention. *Journal of Eye Movement Research*, *1*(2). https://doi.org/10.16910/JEMR.1.2.4

Santangelo, V., & Spence, C. (2007). Multisensory cues capture spatial attention regardless of perceptual load. *Journal of Experimental Psychology: Human Perception and Performance, 33*(6), 1311-1321. https://doi.org/10.1037/0096-1523.33.6.1311

Santangelo, V., Van Der Lubbe, R. H. J., Belardinelli, M. O., & Postma, A. (2006). Spatial attention triggered by unimodal, crossmodal, and bimodal exogenous cues: A comparison of reflexive orienting mechanisms. *Experimental Brain Research*, *173*(1), 40-48. https://doi.org/10.1007/s00221-006-0361-6

Santangelo, V., Ho, C., & Spence, C. (2008). Capturing spatial attention with multisensory cues. *Psychonomic Bulletin and Review*, *15*(2), 398-403. https://doi.org/10.3758/PBR.15.2.398

Schevernels, H., Krebs, R. M., Santens, P., Woldorff, M. G., & Boehler, C. N. (2014). Task preparation processes related to reward prediction precede those related to task-difficulty expectation. *NeuroImage*, *84*, 639-647. https://doi.org/10.1016/j.neuroimage.2013.09.039

Schmitt, M., Postma, A., & De Haan, E. (2000). Interactions between exogenous auditory and visual spatial attention. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 53*(1), 105-130. https://doi.org/10.1080/713755882

Schürmann, M., Grumbt, M., Heide, W., & Verleger, R. (2003). Effects of same- and different-modality cues in a Posner task: Extinction-type, spatial, and non-spatial deficits after right-hemispheric stroke. *Cognitive Brain Research, 16*(3), 348-358. https://doi.org/10.1016/S0926-6410(02)00303-8

Small, D. M., Gitelman, D., Simmons, K., Bloise, S. M., Parrish, T., & Mesulam, M. M. (2005). Monetary incentives enhance processing in brain regions mediating top-down control of attention. *Cerebral Cortex*, *15*(12), 1855-1865. https://doi.org/10.1093/cercor/bhi063

Spence, C., & Driver, J. (1997). Audiovisual links in exogenous covert spatial orienting. *Perception & Psychophysics 1997 59:1*, *59*(1), 1–22. https://doi.org/10.3758/BF03206843

Spence, C., & Driver, J. (1999). A new approach to the design of multimodal warning signals. *Engineering Psychology and Cognitive Ergonomics*, *4*, 455-461.

Spence, C., & Santangelo, V. (2009). Capturing spatial attention with multisensory cues: A review. *Hearing Research*, *258*(1–2), 134–142. https://doi.org/10.1016/j.heares.2009.04.015

Störmer, V. S, McDonald, J. J., & Hillyard, S. A. (2019). Involuntary orienting of attention to sight or sound relies on similar neural biasing mechanisms in early visual processing. *Neuropsychologia*. https://doi.org/10.1016/j.neuropsychologia.2019.107122

Talsma, D., Senkowski, D., & Woldorff, M. G. (2009). Intermodal attention affects the processing of the temporal alignment of audiovisual stimuli. *Experimental Brain Research*. https://doi.org/10.1007/s00221-009-1858-6

Van der Burg, E., Olivers, C. N. L., Bronkhorst, A. W., & Theeuwes, J. (2008). Pip and pop: Nonspatial auditory signals improve spatial visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(5), 1053-1065. https://doi.org/10.1037/0096-1523.34.5.1053

Vroomen, J., & De Gelder, B. (2000). Sound enhances visual perception: Cross-modal effects of auditory organization on vision. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(5), 1583-1590. https://doi.org/10.1037/0096-1523.26.5.1583

Ward, L. M. (1994). Supramodal and modality-specific mechanisms for stimulus-driven shifts of auditory and visual attention. *Canadian Journal of Experimental Psychology*, *48*(2), 242-59. https://doi.org/10.1037/1196-1961.48.2.242

Watson, P., Pearson, D., Most, S. B., Theeuwes, J., Wiers, R. W., & Le Pelley, M. E. (2019). Attentional capture by Pavlovian reward-signalling distractors in visual search persists when rewards are removed. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0226284

Watson, P., Pearson, D., Theeuwes, J., Most, S. B., & Le Pelley, M. E. (2020). Delayed disengagement of attention from distractors signalling reward. *Cognition*. https://doi.org/10.1016/j.cognition.2019.104125

Wittmann, B. C., Schott, B. H., Guderian, S., Frey, J. U., Heinze, H. J., & Düzel, E. (2005). Reward-related fMRI activation of dopaminergic midbrain is associated with enhanced hippocampus-dependent long-term memory formation. *Neuron*. https://doi.org/10.1016/j.neuron.2005.01.010

Wright, R. D., & Ward, L. M. (2008). *Orienting of attention*. Oxford University Press.

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