

Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Research Report

Interaction of spatial attention and the associated reward value of audiovisual objects



Roman Vakhrushev^a and Arezoo Pooresmaeili^{a,b,*}

^a Perception and Cognition Lab, European Neuroscience Institute Goettingen-A Joint Initiative of the University Medical Center Goettingen and the Max-Planck-Society, Goettingen, Germany

^b School of Psychology, University of Southampton, Southampton, United Kingdom

ARTICLE INFO

Article history:

Received 29 November 2023

Reviewed 12 February 2024

Revised 12 February 2024

Accepted 22 March 2024

Action editor Ilona Croy

Published online 22 August 2024

Keywords:

Spatial attention

Reward value

Sensory modality

Audiovisual

ERP

ABSTRACT

Reward value and selective attention both enhance the representation of sensory stimuli at the earliest stages of processing. It is still debated whether and how reward-driven and attentional mechanisms interact to influence perception. Here we ask whether the interaction between reward value and selective attention depends on the sensory modality through which the reward information is conveyed. Human participants first learned the reward value of uni-modal visual and auditory stimuli during a conditioning phase. Subsequently, they performed a target detection task on bimodal stimuli containing a previously rewarded stimulus in one, both, or neither of the modalities. Additionally, participants were required to focus their attention on one side and only report targets on the attended side. Our results showed a strong modulation of visual and auditory event-related potentials (ERPs) by spatial attention. We found no main effect of reward value but importantly we found an interaction effect as the strength of attentional modulation of the ERPs was significantly affected by the reward value. When reward effects were examined separately with respect to each modality, auditory value-driven modulation of attention was found to dominate the ERP effects whereas visual reward value on its own led to no effect, likely due to its interference with the target processing. These results inspire a two-stage model where first the salience of a high reward stimulus is enhanced on a local priority map specific to each sensory modality, and at a second stage reward value and top-down attentional mechanisms are integrated across sensory modalities to affect perception.

© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

* Corresponding author. Present address: School of Psychology, University of Southampton, Southampton, United Kingdom.

E-mail address: a.pooresmaeili@soton.ac.uk (A. Pooresmaeili).

<https://doi.org/10.1016/j.cortex.2024.07.013>

0010-9452/© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Our surrounding environment contains a large amount of information whereas our brain's processing capacity is limited. An important strategy to face this challenge is to prioritize the processing of information that is relevant to our current goals and is associated with the most valuable outcome. However, these aspects, i.e., relevance and reward value, may not always be in the same direction. For instance, when waiting for our friend to pick us up at an intersection, we try to focus our attention on the side where our friend is most likely to arrive. At the same time, we may also notice the sound or sight of approaching vehicles that deliver takeout food, and if hungry we may be even more sensitive to these sources of information compared to the familiar sight of our friend's car. Due to the prevalence of such situations, where environmental stimuli should be processed both based on their relevance as well as their reward value, a large body of literature has sought to investigate the underlying mechanisms of selective attention and reward-driven modulation of perception, as well as the interaction between the two (Anderson, 2016a; Anderson et al., 2021; Carrasco, 2011; Chelazzi et al., 2013; Failing & Theeuwes, 2018; Maunsell, 2004; Pessoa, 2015). However, the exact nature of this interaction and its underlying mechanisms have remained unknown.

Converging evidence from behavioral and neurophysiological studies has shown that reward value modulates sensory perception and its neuronal correlates (Baldassi & Simoncini, 2011; Bayer et al., 2017; Hickey et al., 2010; Hughes et al., 2013; Leo & Noppeney, 2014; Pooresmaeili et al., 2014; San Martín et al., 2016; Serences, 2008). Similarly, selective attention affects behavioral and neural responses to sensory inputs (Desimone & Duncan, 1995). The similarity between reward and attention has raised the question of whether they can be dissociated from each other at all (Maunsell, 2004). Previous studies have provided divergent answers to this question. On the one hand, the majority of past studies have reported a strong interaction between reward and attention (Anderson, 2013; Chelazzi et al., 2013; Failing & Theeuwes, 2018; Le Pelley et al., 2016; Yantis et al., 2012), suggesting a *dependence* between the two. For instance, neurophysiological studies showed similar effects of reward and attention on neuronal responses of area V1, and that reward effects are gated by attention, inspiring the idea of a 'unified selection signal' comprising both factors (Stanisor et al., 2013). Other studies showed a different form of interaction where the strength of attention to locations or features of stimuli was gated by reward (Chelazzi et al., 2013, 2014; Pessoa, 2015). These studies hence point to a bi-directional interaction between attention and reward, each influencing the strength of the other. This view is also in line with recent suggestions that reward and attention jointly influence stimulus representations on an *integrated priority map* (Chelazzi et al., 2014; Failing & Theeuwes, 2018), although it is unknown how such an integration is implemented. In contrast to this idea, yet another set of studies provided evidence for the *independence* of reward and attention, as reward effects were unchanged by attentional load (Baldassi & Simoncini, 2011) or the relevance of reward cue to the task (Garcia-Lazaro et al.,

2018), occurring even when rewarded stimuli were outside of conscious awareness (Lunghi & Pooresmaeili, 2023), and during stages of neural processing distinct from attention (Bayer et al., 2017; Rakhshan et al., 2020; Soltani et al., 2021).

Similar to the approach of the latter studies, one possible way to dissociate the effects of reward and attention is to manipulate each factor orthogonally to the other. For instance, if successful detection of a target among distractors leads to the delivery of a reward, target features that are predictive of higher reward not only signal a better outcome (i.e., reward) but also engage attentional mechanisms more strongly than low reward stimuli. However, incidental reward stimuli that are linked to past rewarding experiences or predict reward magnitude independent of the performance in a task are more likely to have less dependence on task-related attentional processes. Likewise, orthogonal manipulation of reward and attention can be achieved when reward information is signaled through a stimulus feature (Baldassi & Simoncini, 2011; Garcia-Lazaro et al., 2018) or sensory modality (Antono et al., 2022; Pooresmaeili et al., 2014; Vakhrushev et al., 2023) that is distinct from the task-relevant target. Combining these, i.e., using previously rewarded stimuli that are task-irrelevant and are delivered through a different sensory modality, is likely to allow the maximum separation between the reward-related and the attentional prioritization of stimulus processing. In the current study, we try to shed light on the interaction between reward and attention under such conditions.

We draw on previous findings showing that reward associated stimuli from a different sensory modality (audition) can affect perception in vision (Antono et al., 2022; Pooresmaeili et al., 2014), even when cross-modal reward stimuli (auditory sounds) are irrelevant to the task at hand (i.e., visual orientation discrimination). Since there is evidence for the separation of attentional resources across sensory modalities (Alais et al., 2006; Duncan et al., 1997), these findings may indicate that cross-modal reward stimuli affect visual processing independently of visual attention. More recently (Vakhrushev et al., 2023), we showed that behavioral and electrophysiological (EEG) correlates of reward effects differ between stimuli that are in the *same* or *different* sensory modalities as the target (intra-modal and cross-modal, respectively). Whereas intra-modal reward stimuli led to an early suppression of the visual event-related potentials (ERPs), cross-modal rewards boosted the visual ERPs later in time and more persistently. These results further support the idea that the interaction of attentional and reward mechanisms may depend on the sensory modality of the rewarded stimulus with respect to the target of the task (for functional magnetic resonance imaging -fMRI- evidence see Antono et al., 2023). However, in the latter studies, intra-modal and cross-modal conditions not only differed in how reward was cued but also in whether they involved the processing of a unimodal (only visual) or a bimodal (audiovisual) stimulus. Additionally, in these previous studies selective attention was not systematically manipulated, and therefore the assumed interaction between the sensory modality and attention could not be directly tested.

In the current study, we remedy the shortcomings mentioned above and ask whether the behavioral and neural

effects of reward and attention occur independently from each other. Additionally, we ask whether a putative interaction between attention and reward depends on the sensory modality of reward stimuli. To answer these questions, we modified a behavioral paradigm developed by a previous study (Talsma & Woldorff, 2005) where it is possible to independently control both attention and reward. In this paradigm, participants performed a target detection task on bimodal (audiovisual) stimuli. Spatial attention was manipulated in a block-wise manner, where in each block participants were asked to attend either to the left or to the right visual hemifield and report changes that occurred in the attended hemifield. Reward value was manipulated by associating a specific feature of visual (tilt orientation) or auditory (pitch) stimuli with different magnitudes of monetary reward (high-value or no-value), resulting in stimuli where either one, both or neither of the modalities were associated with reward. Importantly, these features were orthogonal to the target detection task and were not predictive of the reward delivery as reward associations were learned in a separate conditioning task prior to the test phase. Concurrently with this task, EEG data was recorded allowing us to inspect the effects of reward and attention on visual as well as auditory ERPs. All procedures and hypotheses (referred to as H1–H7) were preregistered (<https://osf.io/xte4v>).

We expected to find that both reward value (H1) and allocation of attention (H2) enhance the behavioral and electrophysiological indices of sensory perception. Additionally, we predicted that there would be no significant interaction between reward and attention (H3). The latter hypothesis was based on two lines of reasoning: firstly, by maximal orthogonalization of reward and attention we predicted that both factors could independently affect sensory perception. Secondly, the task imposed a minimum load on the attentional control as each trial only contained one stimulus and attention was manipulated across blocks. These characteristics were considered optimal for allowing the reward information to influence perception both on the attended as well as on the unattended side without a significant cost for the system, thereby leading to reward modulations independently of attention. Additionally, we hypothesized that the reward and attention effects and their interaction might depend on the sensory configuration of the reward stimuli, i.e., whether visual, auditory or both visual and auditory stimuli, are associated with reward (H4–H6). Specifically, stimuli containing high reward stimuli in both modalities were predicted to produce the strongest value-driven effects compared to those that contained unrewarded stimuli. Additionally, we expected similar reward-driven and attentional effects on the visual and auditory ERPs (H7). To provide a preview of our results, we found no main effect of reward value (rejecting H1), a strong effect of attention (confirming H2), and an interaction between reward and attention (rejecting H3) where attentional modulation of ERPs was strongest for the high reward stimuli. Additionally, the latter effect showed a dependence on the sensory modality, with the strongest effects observed for the stimuli which contained an auditory high reward stimulus (confirming H6) across visual and auditory ERPs (confirming H7).

2. Methods

2.1. Participants

The sample size was identical to a previous study (Vakhrushev et al., 2023) and was set to $N = 36$, but to compensate for possible dropouts and outliers we recorded data from $N = 42$ participants. Four participants were removed from the analysis due to their low performance during the behavioral task (accuracy $< 60\%$ in three participants during the conditioning) or poor fixation ($> 70\%$ of trials during conditioning contained an eye movement in one participant). Consequently, the data from 38 participants was included in our final analyses (age: 27.2 ± 5.4 years – mean \pm SD; 19 females; 4 left-handed). Subjects were recruited via a local database and had normal or corrected to normal vision. Before the experiment started and after all procedures were explained, participants gave an informed written consent and participated in a practice session. All procedures were approved by “Universitätsmedizin Göttingen” (UMG) under proposal number 15/7/15.

2.2. Stimuli and task

The behavioral paradigm employed during the conditioning (Fig. 1a) and the main task (Fig. 1b) were adopted from a previous study (Talsma & Woldorff, 2005) and involved the detection of infrequent targets in a rapidly presented stream of unimodal (either auditory or visual during the conditioning) or bimodal (audiovisual during the main task) stimuli. All stimuli were lateralized 15° relative to the vertical meridian and 6° below the horizontal meridian, each presented briefly for a duration of 108.3 msec. Auditory stimuli were either a 1050 Hz or a 350 Hz tone with a linear rise and fall of 10 msec and an amplitude of 75 dB which were convolved with head-related transfer function (HRTF) filters to render them colocalized with the visual stimuli (Algazi et al., 2001) and were delivered through in-ear headphones. Visual stimuli were tilted square-wave gratings (size: 6° visual degrees or about 5.8×5.8 cm, spatial frequency: 3 cycles/degree), oriented $\pm 45^\circ$ and modulated between white and gray colors presented on a gray background. Audiovisual stimuli contained auditory and visual stimuli presented synchronously and at the same location producing the impression of one single bimodal object.

Throughout the experiment, the participants' task was to detect target trials, i.e., trials containing a change, by pressing a “Space” button on a keyboard. Target trials were randomly interleaved with the rest of the trials and constituted 20% of all trials. Target trials were similar to standard trials, except that halfway through the presentation of stimuli a transient drop of stimulus intensity occurred, which caused an impression of the stimulus having a stutter (auditory stimuli) or a flicker (visual stimuli). The amount of stimulus intensity drop was determined individually for each participant at the beginning of the experiment and was set to be at a level that targets in each modality could be detected with 90% accuracy. In bimodal stimuli, targets contained a simultaneous change of intensity in both modalities akin to (Talsma & Woldorff, 2005).

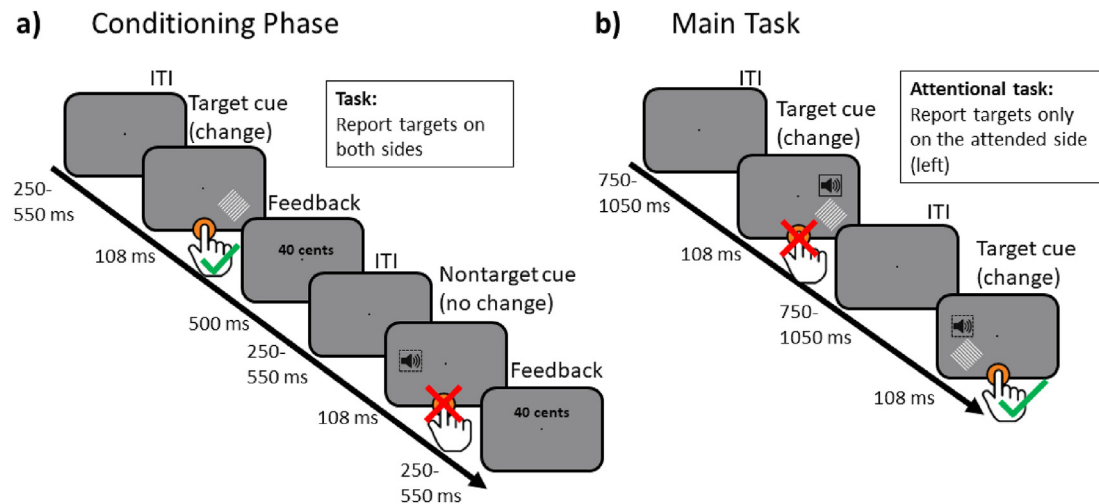


Fig. 1 – Behavioral paradigm. Participants were instructed to detect occasional targets by pressing a “space bar button”. Targets were trials with a change in intensity of auditory and/or visual stimuli midway through their presentation. The change made an impression of a stutter (auditory condition) or a flicker (visual condition). Nontarget (standard) trials contained a stimulus without a change in its intensity and did not require a key press response. Visual ($\pm 45^\circ$ tilt) or auditory stimuli (350 Hz and 1050 Hz) were presented peripherally (15°) and below the fixation point (6°). a) The experiment started with a conditioning phase in which reward associations were learned. In this phase either an auditory or a visual stimulus was presented and participants learned the reward associated with each sensory modality. After every trial, a feedback display showed the reward magnitude that was associated with the presented stimulus. b) In the main task, visual and auditory stimuli always appeared together: unilaterally and synchronously. At the beginning of every block with the main task, participants were instructed to pay attention only to one side of the screen and ignore target trials on the other side (in the demonstration above participants were instructed to only report changes on the left side).

2.3. Conditioning phase

All trials in the conditioning phase contained unimodal (only visual or only auditory) stimuli and were grouped into blocks of 27 trials by the modality. Blocks with different modalities were pseudo-randomly interleaved across the conditioning phase. Participants were instructed to report target trials (trials with a transient change) on both sides of the screen.

During the conditioning phase, participants learned to associate different tones and line orientations with either a positive or zero monetary reward (high-value or no-value, respectively). To achieve this goal, all trials in conditioning phase were followed by a feedback display (500 msec) showing the monetary reward outcome assigned to the presented stimulus. Note that the feedback displays always showed the reward magnitude assigned to a stimulus irrespective of participants' responses. The assigned rewards to the stimuli were drawn from two Poisson distributions with the mean of 40 cents and SD of 5.5 cents for high-reward trials (maximum reward was fixed at 50) and 0 cents for no-reward trials. Participants were instructed to remember the reward outcome of each condition and later report it (at the end of each block in the conditioning phase and main task, through a 2AFC procedure in which the high reward stimulus of two consecutively presented stimuli had to be reported).

In total, the conditioning phase had four conditions: (1) high-value visual, (2) high-value auditory, (3) no-value visual, and (4) no-value auditory (Fig. 2a). These conditions were divided into factors *modality* (visual or auditory) and *reward*

(high- or no-value). Participants completed 960 trials during the conditioning phase, where each condition was repeated 240 times (120 trials per screen side). 20% of all trials were targets (trials with a transient change). To ensure that reward-driven effects were not confounded by effects due to the physical difference between stimuli, two measures were taken. Firstly, the association of visual and auditory stimuli with high or no reward was counterbalanced across participants. Secondly, in each participant, reward associations were reversed halfway through the experiment. To do this, conditioning trials were divided to two halves (480 trials each), one conducted in the beginning and one in the middle of the experiment. The reward assignments were reversed during the second repetition and participants were informed about this through an instruction display before the second repetition of conditioning commenced.

2.4. Main task

At the beginning of each block (180 trials) of the task, participants were instructed to fixate their eyes on the screen center and pay attention to only one side (either left or right, counterbalanced across blocks, Fig. 1b). Participants were required to report targets (i.e., audiovisual stimuli containing a change in intensity in both modalities) only when they were presented on the attended side.

All trials during the main task contained bimodal stimuli. We employed a 2 by 4 design yielding eight conditions which differed in the locus of attention (attended vs unattended) and reward value (reward on one modality, on both or on none).

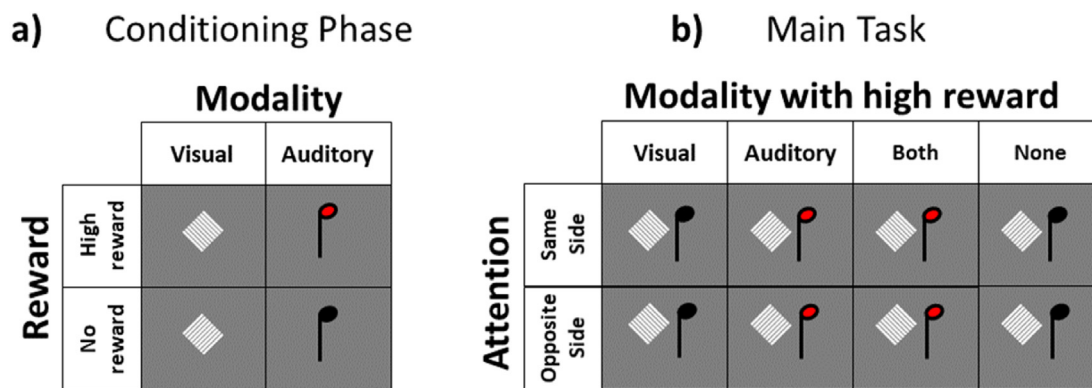


Fig. 2 – Stimuli and design of the experiment. a) Conditioning phase contained two independent factors: stimulus modality (visual or auditory) and reward value (high reward or no reward). b) In the main task two independent factors were manipulated: modality with high reward value (visual, auditory, both, none) and attention (same side or opposite side relative to the audiovisual stimulus).

Specifically, the different reward value configurations consisted of: (1) visual high-value and auditory no-value: VHSN (2) visual no-value and auditory high-value: VNSH (3) visual high-value and auditory high-value: VSHS (4) visual no-value and auditory no-value VNSN (see Fig. 2b). Participants completed 3200 trials during the main task where each of the reward conditions was repeated 400 times (200 trials per screen side). Of these, target trials (trials with a transient change in the intensity of both modalities) constituted 10% of all trials (i.e., 40 trials) on each side (10% on the attended side and 10% on the unattended side). Note that halfway through the experiment, reward associations were reversed. Accordingly, the main task was divided to two repetitions (each 1600 trials), each occurring after the corresponding conditioning phase. This enabled us to measure how the specific reward associations that participants had learned in the preceding conditioning phase influenced the responses in the main task.

In both conditioning and main task, a no-stimulus condition was added with the same number of repetitions as the other conditions (conditioning: 240 trials and main task: 400 trials). In no-stimulus trials only the fixation point was displayed on the screen for the entire duration of the trial without any physical stimulus being presented. This condition was added to eliminate the overlapping electrophysiological activity elicited by the rapidly presented consecutive events (Woldorff, 1993), as implemented by (Talsma & Woldorff, 2005).

2.5. Apparatus

The experiment was conducted in a darkened, sound-attenuated, and electromagnetically shielded chamber. Participants were seated on a chair with their heads fixed on a chinrest positioned 80 cm from a 22.5" monitor (refresh rate = 120 Hz). Stimulus presentation was controlled by a PC under the Windows operating system equipped with MATLAB (version R2015b) and Psychophysics toolbox-3 (Brainard, 1997).

This study assessed the accuracy and reaction times (RT) of the detection of target stimuli (containing a transient change) based on participants responses indicated by a keypress on a keyboard. Eye movements were recorded with an EyeLink 1000 eye tracker system (SR Research, Ontario, Canada) in a

desktop mount configuration, recording the right eye at a sampling rate of 1000 Hz. The EEG data was continuously recorded from 64 electrodes with an actiChamp system and referenced to the A2 electrode. The recording was done with BrainVision software (BrainVision Recorder 1.23.0001 Brain Products GmbH, Gilching, Germany; actiCap, Brain Products GmbH, Gilching, Germany). The signal was digitized at 1000 Hz and amplified with a gain of 10,000. All electrode impedances were kept below 10 k Ω .

2.6. Transformations

2.6.1. Behavioral data

Reaction times (RTs) of correctly detected targets were computed separately for each condition at each phase of the experiment. Mean reaction time was used as the independent variable to assess the effect of reward and attention.

2.6.2. ERP analysis

EEG data was imported and processed offline using EEGLAB (Delorme & Makeig, 2004), an open-source toolbox running under the MATLAB environment. Raw data of each participant was band-pass filtered with .1 Hz as the high-pass cutoff and 40 Hz as the low-pass cutoff frequencies. An automatic bad channel detection algorithm was applied using EEGLAB's pop_rejchan method (threshold = 5, method = kurtosis). All trials detected online as a keypress or fixation breaks (eye position to the fixation point > .9°) were removed from the analysis. After this, epochs were extracted from the remaining data using a stimulus-locked window of 3000 msec (1000–2000 msec) and were subjected to an Independent Component Analysis (ICA) algorithm (Delorme & Makeig, 2004). Eye-blinks and eye-movements artifacts were automatically identified and corrected using an ICA-based automatic method, implemented in the ADJUST plugin of EEGLAB (Mognon et al., 2011). Bad channels were interpolated by using the default spherical interpolation method. Data were re-referenced offline to the average reference. Finally, stimulus-locked epochs were extracted using a window of 1100 msec (–100 to 1000 msec), and baseline corrected using the pre-stimulus time interval (–100 to 0 msec).

Our ERP analysis was focused only on standard trials, i.e., trials without a change, as done previously (Talsma & Woldorff, 2005). Visual ERPs were measured in occipital electrodes (PO7/8 and O1/2) for P1 (70–170 msec) component. Auditory ERPs were measured at Fz for the auditory N1 (90–160 msec) component. P300 (250–400 msec) responses were measured over Pz. Analysis of visual P1 and auditory N1 components was based on the average within 60 msec window around the respective peak of each component measured across participants for each condition. Before the ERP analysis, we removed the overlapping ERP activity from the adjacent trials by subtracting a waveform of a no-stimulus condition from a waveform of each condition with an audio-visual cue, as Talsma & Woldorff (2005) implemented.

Note that based on the visualization of the ERP traces, we changed the following measures compared to the pre-registered plan (see the Supplementary information for results based on the pre-registered plan): firstly, we changed the auditory N1 window from 70–170 msec to 90–160 msec. Secondly, we increased the length of the time window around the peak of visual P1 and auditory N1 components within which ERP amplitudes were averaged from the pre-registered length of 30 msec to 60 msec to account for the large inter-individual difference in the timing of early ERP components. Thirdly, the P300 window was pre-registered to be identical to our previous study (Vakhrushev et al., 2023), i.e., 350–600 msec. However, noting the differences between tasks used in these studies and based on the effects reported by (Talsma & Woldorff, 2005), P300 window was changed to an earlier interval between 250 and 400 msec. Additionally, to be consistent with the method employed by (Talsma & Woldorff, 2005) visual ERPs were inspected in the contralateral electrodes, while we also report the results when ipsilateral visual ERPs were inspected (see the Supplementary information).

As we report in the Supplementary information, all our results were replicated when using the pre-registered settings. However, attentional effects on the auditory N1 responses measured at Fz depended on the choice of time windows (see Figs. 4b and S3b).

2.7. Analysis

2.7.1. Behavioral data

Statistical analyses were done separately for the conditioning phase and the main task. In the conditioning phase, a two-way repeated-measures analysis of variance analysis (ANOVA) of accuracies and RTs was done with factors *modality* (visual or auditory) and *value* (high-value cue or no-value cue, see the Supplementary information).

In the main task (Fig. 3), all stimuli were bimodal, where individual sounds and visual stimuli were either associated with high-value or no-value (VHSH – visual cue has high-value and sound has high-value; VHSN – visual cue has high-value and sound has no-value; VNSH – visual cue has no-value and sound has high-value; VNSN – visual cue has no-value and sound has no-value). Here a two-way ANOVA was used with factors *auditory reward value* (high or no value) and *visual reward value* (high or no value). Consequently, each bimodal stimulus was modeled by both factors: for instance, VHSH was modeled as (*auditory reward value* = high and visual

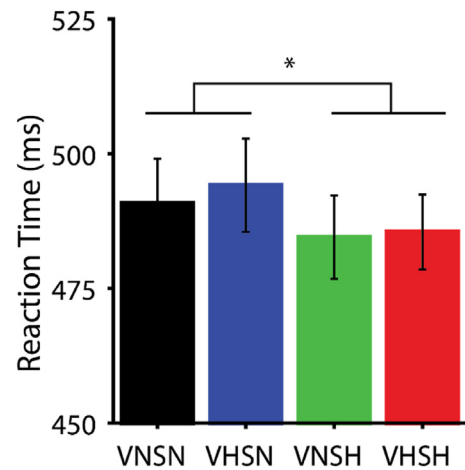


Fig. 3 – Reaction times during the main task. Four stimulus configurations consisted of conditions with high reward value in both modalities (VHSH), only in visual modality (VHSN), only in auditory modality (VNSH), or in neither of the modalities (VNSN). Conditions with reward value in auditory modality (VHSH and VNSH) had shorter reaction times than other conditions, captured by a significant main effect of auditory reward value. Error bars depict s.e.m across participants.

reward value = high), whereas VHSN was modeled as (*auditory reward value* = no and visual reward value = high). Significant main or interaction effects were subsequently examined using planned pairwise comparisons.

2.7.2. EEG data during the main task

Firstly, to test whether we replicate the reported effects of a previous study that had inspired our paradigm (Talsma & Woldorff, 2005), ERP responses to the attended and unattended stimuli across all conditions were compared (Fig. 4).

Subsequently, a rmANOVA comprising factors *attention* (attended or unattended) and *value* (high- or no-value) was done on the data of conditions in which both visual and auditory stimuli of bimodal stimuli had the same value, i.e., either high- or no-value (VHSH and VNSN). Here, a main effect of reward would confirm hypothesis 1, a main effect of attention would confirm hypothesis 2, the absence of interaction effects would confirm hypothesis 3 of our pre-registered plan (Fig. 5).

Having established an effect of attention in previous two analyses, we next subtracted the ERPs of attended and unattended conditions and tested the effect of reward value separately in each modality (Fig. 6). To this end, a two-way rmANOVA with factors *value in visual modality* (high-value or no-value), *value in auditory modality* (high-value or no-value) was done separately on visual P1 and auditory N1 components (*hypothesis 4–6*).

Finally, to test whether the effect of value in visual or auditory modalities differs between different regions of interest (ROI), we conducted a three-way rmANOVA with factors ROI [two levels: visual (PO7, O1, O2, PO8) or frontal region (Fz)], *value in visual modality* (high-value or no-value), *value in auditory modality* (high-value or no-value). Since P1 and N1 components have an opposite amplitude polarity, the amplitude

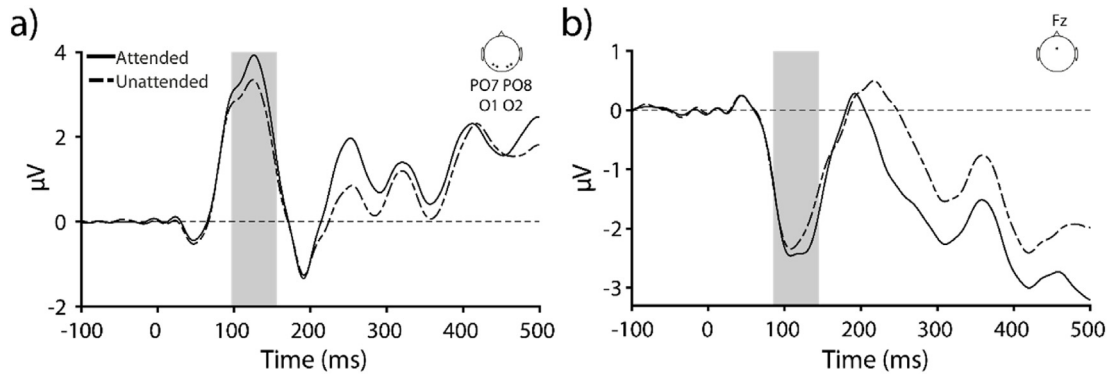


Fig. 4 – Attentional modulation of ERPs. a) ERP waves produced by the attended and unattended audiovisual stimuli that were recorded over contralateral occipital electrodes (PO7, O1, O2, PO8) or b) frontal electrodes (Fz).

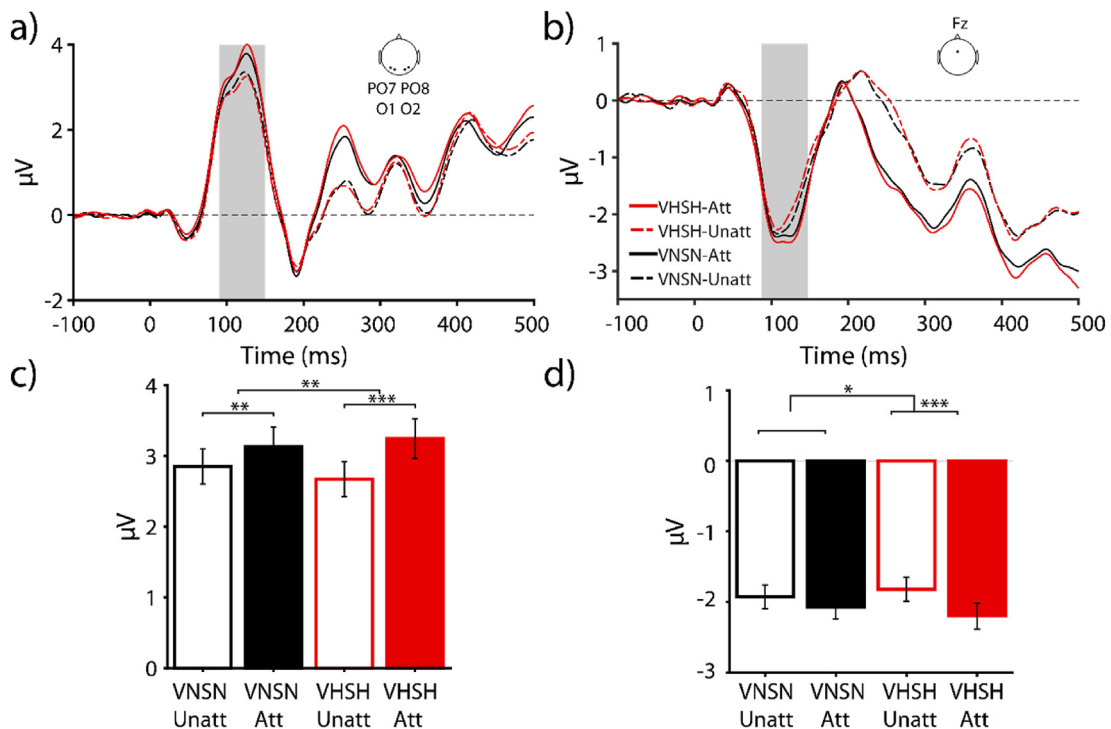


Fig. 5 – Modulation of early ERP components by attention and reward factors. ERP waves measured over contralateral occipital a) or frontal b) regions for conditions that had either high reward value in both modalities (VSHH) or no reward values in both modalities (VNSN) and were presented either on the attended (Att) or on the unattended (Unatt) side. c) Same as a) depicting the average visual P1 amplitudes. d) Same as b) depicting the auditory N1 amplitudes. Error bars depict s.e.m across participants.

of N1 component was rectified (multiplied by -1) before being included in this analysis. Here, our focus was on testing whether an interaction existed between value in each modality and the ROI (*hypothesis 7*).

3. Results

3.1. Behavioral results

The debriefing results and the analysis of behavioral and ERP responses (P300 component) during the conditioning

indicated that all participants had successfully learned the reward associations of both modalities, although the behavioral and ERP effects of reward were stronger in the auditory modality (see the [Supplementary information](#) and [Fig. S1](#)).

During the main task, participants had overall near perfect performance: they correctly reported a target when it was presented on the attended side (hit rate: mean \pm s.e.m. = 95.69% \pm .74) but not when it was on the unattended side (hit rate: mean \pm s.e.m. = .98% \pm .26). Additionally, there was no tendency to erroneously report a target when it was absent (false alarm rates: mean \pm s.e.m. = 1.56% \pm .34 and

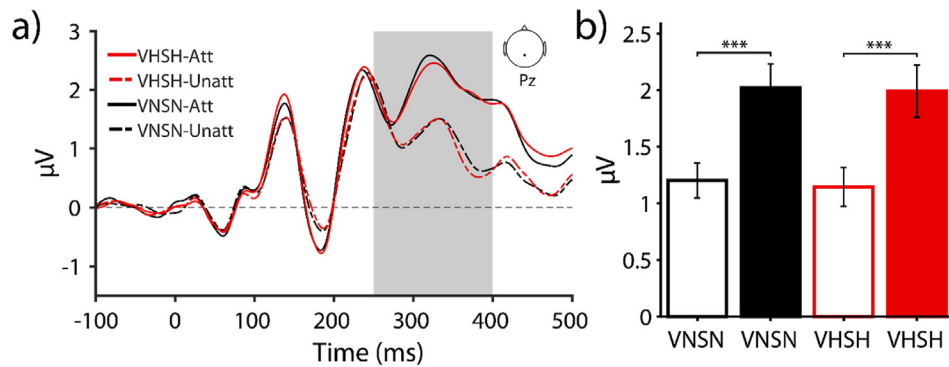


Fig. 6 – Modulation of late ERP components by attention and reward factors. a) ERP waves measured over Pz for conditions that either had high reward value in both modalities (VSH) or no reward values in both modalities (VNSN) and were presented either on the attended (Att) or on the unattended (Unatt) side. **b)** average amplitude of P300 (250–400 msec). Error bars depict s.e.m across participants.

.34% ± .07, on the attended and unattended sides, respectively). To examine how the associated value of visual and auditory stimuli affected the perceived salience of bimodal stimuli, we next analyzed the reaction times (see Fig. 3 and Table 1). Note that reaction times were analyzed for hits on the attended side, as participants had very few errors (i.e., false alarms on either side or hits on the unattended side). A two-way rmANOVA with the factor *visual reward value* (2 levels) and *auditory reward value* (2 levels) revealed only a main effects of *auditory reward value* $F(1,37) = 5.36, p = .026, \eta_p^2 = .127$. This effect corresponded to shorter reaction times in conditions with a high-value auditory stimulus (VHSH and VNSH) compared to other conditions (VHSN and VNSN). Other main and interaction effects did not reach statistical significance ($p > .1$). An exploratory analysis where the phase of the experiment, i.e., before or after the reward assignments were reversed, was entered as an additional factor in the above ANOVA only confirmed a significant effect of *auditory reward value* with neither a main effect nor a significant interaction of phase with reward value ($p > .1$).

Together, the analysis of behavioral data in conditioning and main task indicates a robust influence of reward value on target detection in the auditory modality.

3.2. Attentional modulation of ERP responses

We first tested whether the attentional modulation effects reported by Talsma et al. (2005) are replicated in our paradigm, as shown in Fig. 4. A one-way ANOVA on visual P1 component (PO7, O1, O2, PO8) revealed a main effect of *attention* $[F(1,37) = 41.17, p < .001, \eta_p^2 = .53]$. Attended conditions

increased the amplitude of P1 component (mean ± SD: $3.20 \pm .27$) compared to the unattended conditions (mean ± SD: $2.78 \pm .24$). A similar effect was found in auditory N1 component measured at Fz $[F(1,37) = 15.53, p < .001, \eta_p^2 = .30]$, where attended stimuli increased the ERP amplitude of N1 component (mean ± SD: $-2.13 \pm .17$) compared to unattended stimuli (mean ± SD: $-1.90 \pm .15$). These results replicate previous findings (Talsma & Woldorff, 2005) and indicate that our experimental design modulated the allocation of spatial attention successfully.

3.3. The interplay of attention and reward value in modulating ERP responses

3.3.1. Visual P1 component

Next, we examined the attentional and reward effects in conditions where both modalities had either high- or no-value (VHSH and VNSN, see Fig. 5a,c). Here, an rmANOVA (factors *attention*: attended/unattended and *reward value*: high/no) on the visual P1 component of the ERPs revealed a main effect of attention $[F(1,37) = 24.08, p < .001, \eta_p^2 = .39]$ but no main effect of reward value $[F(1,37) = .13, p = .71]$. Attended stimuli evoked higher amplitudes of P1 component (mean ± s.e.m.: $3.19 \pm .27$), compared to unattended stimuli (mean ± s.e.m.: $2.76 \pm .24$). Interestingly, the difference between attended and unattended conditions was larger for high-value (mean ± s.e.m.: $.58 \pm .10, t(37) = 5.73, p < .001, d_z = .93$) compared to no-value (mean ± s.e.m.: $.28 \pm .10, t(37) = 2.79, p = .008, d_z = .45$) stimuli, leading to a significant interaction effect between attention and reward $[F(1,37) = 8.38, p = .006, \eta_p^2 = .18]$.

Table 1 – Overview of behavioral and electrophysiological indices measured during the main task in each stimulus condition.

	VNSN	VHSN	VNSH	VHSH
Response time (RT, ms)	490.79 ± 8.28	494.14 ± 8.65	484.50 ± 7.73	485.47 ± 6.93
Hit rate (%)	95.20 ± .87	95.79 ± .77	95.59 ± .83	96.18 ± .83
Visual P1 amplitude (attended condition)	3.14 ± .27	3.22 ± .27	3.21 ± .26	3.25 ± .28
Visual P1 amplitude (unattended condition)	2.85 ± .24	2.86 ± .24	2.73 ± .25	2.63 ± .25
Auditory N1 amplitude (attended condition)	-2.07 ± .18	-2.13 ± .18	-2.14 ± .17	-2.20 ± .19
Auditory N1 amplitude (unattended condition)	-1.93 ± .17	-1.98 ± .15	-1.89 ± .16	-1.82 ± .17

3.3.2. Auditory N1 component

Similar to the P1 component, an analysis of N1 component (Fig. 5b, d) revealed a main effect of attention [$F(1,37) = 9.2$, $p = .004$, $\eta_p^2 = .2$], no effect of reward value [$F(1,37) = .01$, $p = .91$], and an interaction effect between attention and reward value [$F(1,37) = 6.10$, $p = .018$, $\eta_p^2 = .142$]. Attended stimuli evoked a more negative N1 peak (mean \pm s.e.m.: $-2.13 \pm .18$) compared to unattended stimuli (mean \pm s.e.m.: $-1.87 \pm .16$). Additionally, high-value stimuli resulted in a larger modulation of N1 amplitude by attention (mean \pm s.e.m.: $-.38 \pm .09$, $t(37) = 4.1$, $p < .001$, $d_z = .66$) compared to no-value stimuli (mean \pm s.e.m.: $-.14 \pm .10$, $t(37) = 1.34$, $p = .18$, $d_z = .21$). The difference between attended and unattended no-value stimuli did not reach significance ($p = .188$).

3.3.3. P300 component

We next examined the P300 component measured in Pz electrode (250–400 msec, see Fig. 6). We found a strong main effect of attention [$F(1,37) = 56.43$, $p < .001$, $\eta_p^2 = .60$] but neither a main effect of reward value [$F(1,37) = .67$, $p = .41$] nor an interaction between reward value and attention [$F(1,37) = .00$, $p = .96$]. An exploratory analysis of P300 component in Fz electrode (visible in Fig. 5) showed a strong attentional effect [$F(1,37) = 37.68$, $p < .001$, $\eta_p^2 = .51$] but no main effect of reward [$F(1,37) = .20$, $p = .65$], and only a non-significant numerical trend for an interaction between attention and reward [$F(1,37) = 3.49$, $p = .07$, $\eta_p^2 = .086$].

Together, the results obtained while examining the early visual and auditory ERP components (visual P1 and auditory N1) showed a main effect of attention and an interaction

between attention and reward value. A stronger attentional modulation for high-value stimuli demonstrates that stimulus value influences the allocation of attention.

3.4. Dependence of the reward-driven modulation of attention on the sensory modality of reward associated stimuli

We have so far shown a strong effect of attention on early ERP components that is also influenced by the stimulus reward value. We next asked whether this effect depends on the specific configuration of stimuli; i.e., the modality through which reward value is cued (Fig. 7). To this end, the ERPs of attended and unattended conditions were subtracted from each other and entered into an rmANOVA with factors visual reward value and auditory reward value.

Analysis of visual P1 component revealed no main effect of visual reward value [$F(1,37) = 1.61$, $p = .21$, $\eta_p^2 = .04$] but a strong main effect of auditory reward value [$F(1,37) = 9.74$, $p = .003$, $\eta_p^2 = .21$]. No interaction effect was found between these factors [$F(1,37) = .02$, $p = .82$]. Importantly, the modulation of visual cortex by reward value was strongly lateralized and occurred predominantly in the contralateral sites with respect to the stimulus (see Supplementary Fig. S2).

Analysis of auditory N1 component showed a main effect of auditory reward value [$F(1,37) = 5.64$, $p = .023$, $\eta_p^2 = .13$] as the presence of high-value auditory stimuli increased the N1 amplitude (mean \pm s.e.m.: $-.32 \pm .05 \mu\text{V}$) compared to conditions that contained a no-value auditory cue (mean \pm s.e.m.: $-.14 \pm .08 \mu\text{V}$). The main effect of visual reward value

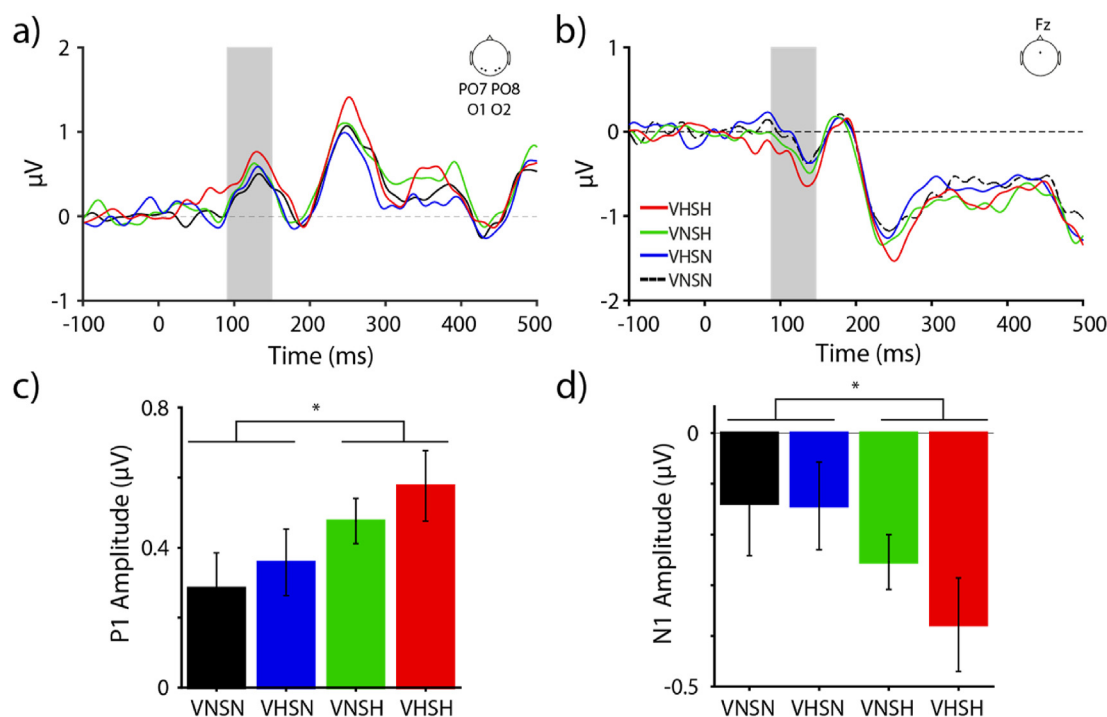


Fig. 7 – Modulation of attention by reward value of auditory and visual modalities. a) The difference wave between attended and unattended reward conditions (high reward in vision: VHSN, in auditory: VNSH, in both: VHSH, or in neither: VNSN modality) measured in contralateral occipital electrodes (PO7, O1, O2, PO8). b) Same as (a) for Fz electrode. c) Same as a) depicting the average visual P1 amplitudes. d) Same as b) depicting the auditory N1 amplitudes. Error bars depict s.e.m across participants.

[$F(1,37) = .94, p = .34$] and the interaction of visual and auditory reward value were not significant [$F(1,37) = .57, p = .45$].

Inspection of individual stimulus configurations (see [Table 1](#) and [Fig. 7](#)) reveals that both P1 and N1 had the highest amplitude when visual and auditory stimuli were associated with high-value (VHSH), followed by high-value only in auditory modality (VNSH) whereas other configurations (VHSN and VNSH) elicited lower amplitudes. This may indicate that when reward value is high in both modalities, the salience of a bimodal stimulus is enhanced in an additive manner. However, we note that this trend was not robust enough to manifest as a significant interaction effect, neither for visual P1 nor for auditory N1 components.

To examine whether the change in reward associations halfway through the experiment impacted our reported results, we conducted an exploratory analysis, entering the *phase* (before or after the change in reward assignments) into the analysis of the visual P1 and auditory N1 ERP data shown in [Fig. 7](#). This exploratory analysis confirmed the aforementioned results, i.e., a significant effect of *auditory reward value*, and additionally only revealed a weak interaction of *phase* with *auditory reward value* for auditory N1 component [$F(1,37) = 4.51, p = .04, \eta_p^2 = .11$]. Therefore, the change in reward associations did not have a strong influence on our reported results that were averaged across phases, as intended.

Finally, to test whether the influence of auditory and visual reward value depended on the region of interest (ROI), we conducted a *rmANOVA* on the visual P1 and auditory N1 components (factors auditory reward value: high/no, visual reward value: high/no, ROI: occipital/frontal). This analysis revealed a main effect of ROI [$F(1,37) = 10.06, p < .001, \eta_p^2 = .21$] as expected, and a main effect of *auditory reward value* [$F(1,37) = 10.01, p < .005, \eta_p^2 = .21$]. No interaction was found between the auditory or visual reward value and ROI [$F(1,37) < 1, p > .1$], indicating that auditory reward value enhanced sensory processing both intra-modally (i.e., in Fz) as well as cross-modally (in occipital ROI).

3.5. Examination of potential confounds: effects of differing discriminability in visual and auditory stimuli

In our experiments, the visual stimuli consisted of relatively high spatial frequency gratings presented peripherally (a grating with a size of 6° and a spatial frequency of 3 cycles/degree presented at 15° eccentricity). Given the potential difficulty in detecting these visual targets and distinguishing their orientations, participants might have ignored visual cues during the main task, relying solely on auditory cues. Since changes in stimulus intensity occurred in both modalities in every trial during the main task, this might have been a plausible strategy for participants. If this were the case, we would observe a significant difference between the detection accuracy of visual and auditory targets during the conditioning phase, when these targets were presented alone (unimodally), with visual targets being detected with significantly lower accuracy. While the former analysis examines the difference in the detection of visual and auditory stimuli, we further tested whether the distinction between reward cues in the visual ($\pm 45^\circ$ tilt orientations of gratings) or auditory modality (1050 Hz or 350 Hz) was significantly different and

whether the degree to which these cues were discriminable influenced our observed effects (behavioral and ERP) during the main task. To test these possibilities, we conducted the following analyses.

First, we analyzed the detection accuracies of visual and auditory targets during the conditioning phase (see also the [Supplementary information](#)). We found no main effect of modality, indicating that visual targets were not detected with significantly lower accuracy than auditory targets (accuracies were mean \pm s.e.m. = $\%89.91 \pm 1.36, \%87.32 \pm 1.33, \%91.39 \pm 1.38, \text{ and } \%84.70 \pm 1.79$ for Visual High Reward: VH, Visual No Reward: VN, Auditory High Reward: SH, and Auditory No Reward: SN configurations, respectively). These detection accuracies were averaged across the two halves of the experiment, i.e., before and after the reward associations switched. To rule out the possibility that the detection of visual or auditory targets may have differed across different phases of the experiment, we then examined each phase separately. Two repeated measures ANOVAs were conducted with the accuracy of target detection during the conditioning task as the dependent variable, and reward value and modality as the independent variables. These analyses revealed a significant main effect of reward value [$F(1,37) = 18.5, p = .0001$ before the switch; $F(1,37) = 5.10, p = .03$ after the switch] and no main or interaction effect with modality [all $F_s < 1$ and $p_s > .1$, except for a trend for an interaction effect between reward and modality after the switch with $F(1,37) = 3.26, p = .08$]. These analyses thus rule out the possibility that participants were unable to detect the visual stimuli or their reward associations in either phase of the experiment. Additionally, our debriefing results indicated that all participants could correctly report the reward value in both vision and audition (see the [Supplementary information](#)).

Secondly, to rule out the possibility that the behavioral effects observed during the main task were due to the numerical difference in the discriminability of the reward associations of visual compared to auditory cues, we conducted two exploratory correlation analyses for each phase of the experiment. Pearson correlations between reward effects on reaction times (RTs) during the main task (VHSN–VNSH) and the discrimination accuracy of reward cues in each modality [(VH–VN) – (SH–SN)] revealed no significant relationship in either phase (all $p_s > .1$). These results indicate that the weaker reward effects in the visual modality (in VHSN compared to VNSH) during the main task were not due to participants' difficulty in discriminating the reward value of visual compared to auditory cues.

While the above analyses do not suggest that participants could not detect the visual stimuli or discriminate their reward associations, we next explicitly modeled the contribution of any difference in discriminability of reward effects in the two modalities and re-examined the behavioral effects depicted in [Fig. 3](#) and P1 ERP effects depicted in [Fig. 7](#). Specifically, we included the difference in cue discriminability across the visual and auditory modalities during the conditioning task [(VH–VN) – (SH–SN)] as a between-subject covariate in our ANOVAs. The analysis of reaction times during the main task ([Fig. 3](#)) again showed a main effect of auditory reward $F(1,36) = 5.38, p = .026, \eta_p^2 = .13$ and only a numerical trend for reward effects in the visual modality $F(1,36) = 3.30,$

$p = .077$, $\eta^2 = .084$ without any interaction with the between-subject covariate. Likewise, the analysis of P1 responses (Fig. 7a and c) while including difference in cue discriminability as a between-subject covariate reproduced our previous results with a significant effect of auditory reward value $F(1,36) = 7.33$, $p = .01$, $\eta^2 = .17$. Additionally, numerical trends for an effect of visual reward $F(1,37) = 3.66$, $p = .064$, $\eta^2 = .09$ and an interaction between visual reward and cue discriminability $F(1,37) = 3.93$, $p = .055$ were observed, but these effects did not reach significance. We limited this exploratory analysis to the P1 component as the discriminability of visual stimuli is most likely to affect this component, if at all. Therefore, considering the effect of cue discriminability led to a slight increase in the effect size of visual reward (cf. η^2 of this factor with and without including the discriminability as a covariate). However, the contribution of cue discriminability remained a non-significant numerical trend.

Together, these analyses rule out the possibility that participants had significantly lower detection of visual compared to auditory stimuli or were unable to discriminate the reward value of visual cues.

4. Discussion

The current study examined the interaction between spatial attention and reward value when visual, auditory, or both modalities were associated with high reward. During the conditioning phase, both visual and auditory high reward stimuli facilitated target detection, with a stronger effect in the auditory modality. During the main task, we found a strong effect of attention on early event-related responses of occipital and frontal areas (visual P1 and auditory N1 components, respectively), thus replicating a previous study that inspired our design (Talsma & Woldorff, 2005). We did not find a main effect of reward value, but importantly we found a significant interaction as attentional modulation of P1 and N1 responses were stronger when both modalities had high-value compared to no-value. These findings suggest that reward value modulates the strength of attentional filtering. However, when visual and auditory reward effects were compared separately, a more robust reward modulation of attention was found for the latter both behaviorally and electrophysiologically, thus reflecting the dominant role of auditory reward signals in the guidance of attention. Hence, under the settings employed in our study, stimuli with high value in the auditory domain guide the attentional resources towards the attended location and withdraw them from the unattended location.

A host of previous studies have demonstrated that reward biases selective attention towards stimuli previously associated with a positive outcome, an effect observed for shapes (Della Libera & Chelazzi, 2006, 2009), faces (Raymond & O'Brien, 2009), orientations (Kiss et al., 2009), and semantic information in Stroop tasks (Krebs et al., 2010). However, it is not always possible to determine the degree to which the observed behavioral or neural modulations are due to attention, reward, or both (Maunsell, 2004). Often reward-associated stimuli appear as the target of the task or contain target-related information, hence inadvertently engaging attention. More recent studies showed that even when reward stimuli are task-

irrelevant (Garcia-Lazaro et al., 2018) or are delivered through a different sensory modality (Antono et al., 2022; Pooresmaeili et al., 2014), they still influence target processing, suggesting independence of reward and attentional mechanisms. Specifically, a lack of influence of attentional load on reward effects (Baldassi & Simoncini, 2011), temporally distinct ERP modulations for reward and attention (Bayer et al., 2017), and separable effects of reward on sensory processing and choice (Rakhshan et al., 2020; Soltani et al., 2021) supported the idea that selective attention and reward can influence perception independently. Here, we tried to resolve this ambiguity but while we had predicted the independence of attention and reward, we found evidence to the contrary. In contrast to the previous studies that showed a clear separation between effects of reward and attention, our results show that reward effects on early perception occur exclusively through the modulation of attention. We note several major differences between the latter studies and ours. Firstly, Baldassi et al. (Baldassi & Simoncini, 2011) employed a weak manipulation of spatial attention, i.e., by performing a secondary task at fixation, compared to ours where attention was strictly controlled in space (to be on one or the other side). The form of attentional manipulation by Baldassi and Simoncini may have allowed a boost of reward effects through attention, despite the conclusion of the study (Baldassi & Simoncini, 2011). Secondly, in studies that found separable correlates of reward and attention reward stimuli were predictive of the delivery of reward during the main task (Baldassi & Simoncini, 2011; Bayer et al., 2017; Rakhshan et al., 2020; Soltani et al., 2021), whereas we tested stimuli that were previously associated with reward value and did not lead to rewards anymore. Reward predicting incentives can overall exert stronger effects on perception as better perceptual performance is constantly reinforced by the delivery of reward (Antono et al., 2022). This however does not allow to study the pure role of the associated reward value independent of any task-related feedback or continuous reward reinforcement. In the light of these studies and the fact that in our paradigm reward effects only manifested as an interaction with attention, we surmise that independent effects of attention and reward can only be observed under conditions where reward can exert its maximal influence on perception without any assistance from attentional mechanisms. In other contexts, such as ours where rewards are learned prior to the task and are not further reinforced, the associated reward value of stimuli may affect perception primarily through the modulation of attention.

Another possible scenario is that our protocol placed a greater emphasis on attentional control rather than reward processing. For example, during conditioning, participants were not instructed to regulate their focus of attention; they were merely exposed to reward associations of the stimuli. This stands in stark contrast to the main task, where the primary task requirement was to actively control the focus of attention while disregarding reward-related information. Furthermore, instead of a fine-grained control of the deployment of attention through changing the validity of attentional cueing across trials, we used a paradigm during the main task that required participants to only pay attention to one side over a long block of trials. This feature may have strongly discouraged the processing of reward information on the

unattended side that had been observed in previous studies. However, we note that previous studies have found an effect of reward cues presented at the unattended locations even when the location of the target was known in advance and remained the same on every trial during visual search tasks (Munneke et al., 2016; Wang et al., 2014, 2018). Systematically controlling the attentional and reward-related demands of various tasks while exploring the interdependence or independence of these two factors would be a promising direction for future research.

While early ERP components (visual P1 and auditory N1) exhibited an effect of reward on attentional processing, we did not find any reward-driven modulations in P300 responses during the main task. We believe that two factors may account for this finding. Firstly, P300 modulation underlies the evaluation of different reward outcomes (Wu & Zhou, 2009; Yeung & Sanfey, 2004) and usually occurs only in response to cues that are predictive of the delivery of rewards and less so when reward delivery is halted (Rossi et al., 2017; Sommer & Schweinberger, 1992; Vakhrushev et al., 2023). In fact, during conditioning, when stimuli were predictive of the delivery of different rewards, we found a modulation of P300 responses, especially for auditory stimuli (see Supplementary Fig. S1). Secondly, the task we employed involved detecting infrequent targets in a rapidly presented stream. In such tasks, target identification and processing are reflected in P300 responses (Luck et al., 2000; Polich, 2007). However, our ERP analysis was restricted to the responses to the standard stimuli (nontarget stimuli without a change in intensity), and the processing of these stimuli relies to a lesser extent on P300 responses, as shown before (Brookhuis et al., 1983; Talsma & Woldorff, 2005). This may account for a lack of reward effect on P300 responses during the main task, despite the robust attentional modulations of this component.

An important novel aspect of the current study is to investigate the effects of visual and auditory stimuli that were previously associated with high monetary reward outcomes. Previous studies on value-driven mechanisms have been mostly focused on reward effects in the visual modality (for a review see Failing & Theeuwes, 2018), although reward effects have been reported in all sensory modalities (Goltstein et al., 2013; Pleger et al., 2008; Rutkowski & Weinberger, 2005; Shuler & Bear, 2006; Stanisor et al., 2013; Weil et al., 2010). This leaves a question of the extent to which value-driven mechanisms reflect a general principle of information processing across sensory modalities. A few recent studies approached this question by testing cross-modal reward modulations (Anderson, 2016b; Pooresmaeili et al., 2014; Sanz et al., 2018), testing the competition of reward and attention in different sensory modalities (Anderson et al., 2011a; Asutay & Västfjäll, 2016; Baines et al., 2011; Bourgeois et al., 2016; Chelazzi et al., 2014; Failing et al., 2015; Kang et al., 2017; Kim et al., 2021; Qin et al., 2021; Tankelevitch et al., 2020; Yantis et al., 2012), and comparing intra- and cross-modal reward modulations directly (Anderson, 2016b; Antono et al., 2022, 2023; Bean et al., 2021; Cheng et al., 2020; Kang et al., 2017; Vakhrushev et al., 2023). The emerging picture from these studies is that value-driven mechanisms reflect a general information processing principle that operates across modalities.

One important principle across sensory modalities is that when task-irrelevant features are associated with high reward they may potentially interfere with the processing of the target (Anderson et al., 2021). In line with this, recent studies (Vakhrushev et al., 2023) showed an early suppression of visual ERPs when the task-irrelevant reward cue was in the visual modality (i.e., intra-modally), but surprisingly late ERP components and behavioral sensitivity were enhanced when reward was cued through the auditory modality (Pooresmaeili et al., 2014; Vakhrushev et al., 2023). Hence, reward information exerted different effects on sensory processing dependent on whether it was cued intra- or cross-modally suggesting that modality-specific attentional resources might contribute to value-driven effects (see also Antono et al., 2023; Hoofs et al., 2022). Interestingly, although we used a different task and explicitly controlled the amount of spatial attention towards reward stimuli, our results are in line with the interpretation provided by (Vakhrushev et al., 2023). While visual high-value stimuli on their own (i.e., in VHSN configuration) led to no reward modulation of visual P1, the joint presence of high value in both modalities (i.e., VSHS) produced the strongest reward-driven modulation of attention (Fig. 7). These two findings reveal two important stages of reward-driven modulation of sensory processing. The *first stage* occurs locally within each sensory modality. At this stage, reward information competes with attention in order to gain priority in sensory processing (Anderson et al., 2011b, 2021; Failing & Theeuwes, 2018). As reward information conveyed through vision could potentially compete with the target detection task, for instance by withdrawing resources from a relevant feature (luminance) and allocating them to an irrelevant feature (line orientation), the putative local competition of reward and attention can lead to a reduction of reward effects. However, the local competition can be overturned by rewards in another sensory modality, suggesting that at a *second stage*, the reward information is integrated across different sensory modalities (Bruns et al., 2014; Cheng et al., 2020) resulting in a general improvement of sensory representations. Future studies will be needed to test the validity of this proposal.

Our results demonstrate a distinct pattern of interaction between reward information and the allocation of spatial attention in the visual compared to the auditory modality. This difference can be due to differences in the underlying mechanisms of reward processing in the two sensory modalities, but it also can be related to the specific features of our experiment. For instance, the type of the task that we used, where a change in a rapidly presented stream of audiovisual stimuli had to be detected, may have been easier to be performed through the auditory modality, as visual targets were presented peripherally and were potentially more difficult to detect. However, our analyses revealed that the detection accuracy of visual and auditory cues were not significantly different when cues of each sensory modality were presented alone during the conditioning, and hence participants should have been able to detect both visual and auditory cues when they were presented together. Another possibility is that participants could not discriminate the reward associations of visual cues as much as they were able to tell different auditory cues apart. We checked this potential confound and found no significant contribution

of a difference in cue discriminability on the behavioral and P1 ERP responses. These analyses rule out the possibility that our effects are due to participants' inability to detect visual cues or discriminate their reward associations. Nonetheless, while we decided to use a design similar to the original study (Talsma & Woldorff, 2005) where on each trial intensity changes corresponding to targets occurred in both visual and auditory streams, an alternative design in which targets could alternate between modalities would be more suitable for future studies and can ensure that both modalities are equally used for performing the task. Additionally, and as mentioned before, different attentional requirements during the two phases of the experiment (reporting the target on either side during conditioning vs only on the attended side during the main task), may have resulted in a further downregulation of the visual reward cues, which could potentially interfere with the allocation of attention during the main task. Furthermore, our study may have been underpowered for detecting the interaction effects of visual rewards and attention. These alternatives warrant the usage of larger sample sizes in future studies, a tighter control of experimental factors, and comparable manipulation of attention during the reward associative learning and test phase.

Finally, in our experiment, participants learned the reward value of each sensory modality separately and they were then exposed to bimodal stimuli containing the same or different reward values in two modalities. The pattern of results observed here may be due to this training protocol, as auditory and visual stimuli were never paired during the learning. Alternatively, learning the associated reward value of a bimodal stimulus may promote integrating sensory features and reward values more strongly across sensory modalities and thereby lead to reward modulations that have less dependence on modality-specific mechanisms. Future studies will be needed to examine the role of learning protocols in how reward signals from different sensory modalities interact.

5. Conclusion

In summary, we found that the allocation of spatial attention towards audiovisual stimuli is guided by the associated reward value of auditory and visual modalities. In the context of the task employed here, value-driven modulation of attention was more robust in auditory modality, with the maximum effect observed when both visual and auditory components of an audiovisual stimulus were associated with high value. These results inspire a two-stage model in which reward information is first represented separately in each sensory modality and is subsequently integrated across modalities. The integration of reward value boosts the combined value of a bimodal stimulus and at the same time enhances the attentional selection of the task-relevant information.

Transparency statement

We report how we determined our sample size, all data exclusions, all inclusion/exclusion criteria, whether inclusion/

exclusion criteria were established prior to data analysis, all manipulations, and all measures in the study.

Study material and data availability statement

The processed data and analysis code for each figure are publicly available at <https://osf.io/ajfm4>. Additionally, the raw data and the complete set of study materials can be found at <https://osf.io/h8msu/>.

Open practices

The study in this article has earned Open Data, Open Materials and Preregistered badges for transparent practices. The data, materials and preregistered studies are available at: <https://osf.io/ajfm4>.

CRedit authorship contribution statement

Roman Vakhrushev: Writing – review & editing, Writing – original draft, Visualization, Validation, Project administration, Methodology, Investigation, Formal analysis, Data curation. **Arezo Pooresmaeili:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare no competing interests.

Acknowledgments

We thank Franziska Ehbrecht for her help with the data collection. This work was supported by an ERC Starting Grant (no: 716846) to AP.

Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.cortex.2024.07.013>.

REFERENCES

- Alais, D., Morrone, C., & Burr, D. (2006). Separate attentional resources for vision and audition. *Proceedings. Biological Sciences*, 273(1592), 1339–1345. <https://doi.org/10.1098/rspb.2005.3420>
- Algazi, U., Duda, R., Thompson, D., & Avendano, C. (2001). The CIPIC HRTF database. In *Applications of signal processing to audio and acoustics. IEEE Workshop. on the*, 99–102.
- Anderson, B. A. (2013). A value-driven mechanism of attentional selection. *Journal of Vision*, 13(3). <https://doi.org/10.1167/13.3.7>

- Anderson, B. A. (2016a). 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. (2016b). Value-driven attentional capture in the auditory domain. *Attention, Perception, and Psychophysics*, 78(1), 242–250. <https://doi.org/10.3758/s13414-015-1001-7>
- Anderson, B. A., Kim, H., Kim, A. J., Liao, M.-R., Mrkonja, L., Clement, A., & Grégoire, L. (2021). The past, present, and future of selection history. *Neuroscience and Biobehavioral Reviews*, 130, 326–350. <https://doi.org/10.1016/j.neubiorev.2021.09.004>
- Anderson, B. A., Laurent, P. A., & Yantis, S. (2011a). Learned value magnifies salience-based attentional capture. *PLoS One*, 6(11). <https://doi.org/10.1371/journal.pone.0027926>
- Anderson, B. A., Laurent, P. A., & Yantis, S. (2011b). Value-driven attentional capture. *Proceedings of the National Academy of Sciences of the United States of America*, 108(25), 10367–10371. <https://doi.org/10.1073/pnas.1104047108>
- Antono, J. E., Dang, S., Auksztulewicz, R., & Pooresmaeili, A. (2023). Distinct patterns of connectivity between brain regions underlie the intra-modal and cross-modal value-driven modulations of the visual cortex. *Journal of Neuroscience*, 43(44), 7361–7375. <https://doi.org/10.1523/JNEUROSCI.0355-23.2023>
- Antono, J. E., Vakhrushev, R., & Pooresmaeili, A. (2022). Value-driven modulation of visual perception by visual and auditory reward cues: The role of performance-contingent delivery of reward. *Frontiers in Human Neuroscience*, 16. <https://doi.org/10.3389/fnhum.2022.1062168>
- Asutay, E., & Västfjäll, D. (2016). Auditory attentional selection is biased by reward cues. *Scientific Reports*, 6, Article 36989. <https://doi.org/10.1038/srep36989>
- Baines, S., Ruz, M., Rao, A., Denison, R., & Nobre, A. C. (2011). Modulation of neural activity by motivational and spatial biases. *Neuropsychologia*, 49(9), 2489–2497. <https://doi.org/10.1016/j.neuropsychologia.2011.04.029>
- Baldassi, S., & Simoncini, C. (2011). Reward sharpens orientation coding independently of attention. *Frontiers in Neuroscience*. <https://doi.org/10.3389/fnins.2011.00013>. FEB.
- Bayer, M., Rossi, V., Vanlessen, N., Grass, A., Schacht, A., & Pourtois, G. (2017). Independent effects of motivation and spatial attention in the human visual cortex. *Social Cognitive and Affective Neuroscience*, 12(1), 146–156. <https://doi.org/10.1093/scan/nsw162>
- Bean, N. L., Stein, B. E., & Rowland, B. A. (2021). Stimulus value gates multisensory integration. *European Journal of Neuroscience*. <https://doi.org/10.1111/ejn.15167>
- Bourgeois, A., Chelazzi, L., & Vuilleumier, P. (2016). How motivation and reward learning modulate selective attention. In *Progress in brain research* (Vol. 229, pp. 325–342). Elsevier B.V.. <https://doi.org/10.1016/bs.pbr.2016.06.004>
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10(4), 433–436. <https://doi.org/10.1163/156856897X00357>
- Brookhuis, K. A., Mulder, G., Mulder, L. J. M., & Gloerich, A. B. M. (1983). The P3 complex as an index of information processing: The effects of response probability. *Biological Psychology*, 17(4), 277–296. [https://doi.org/10.1016/0301-0511\(83\)90004-2](https://doi.org/10.1016/0301-0511(83)90004-2)
- Bruns, P., Maiworm, M., & Röder, B. (2014). Reward expectation influences audiovisual spatial integration. *Attention, Perception, and Psychophysics*, 76(6), 1815–1827. <https://doi.org/10.3758/s13414-014-0699-y>
- Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research*, 51(13), 1484–1525. <https://doi.org/10.1016/J.VISRES.2011.04.012>
- Chelazzi, L., Eštočinová, J., Calletti, R., Gerfo, E. L., Sani, I., Libera, C. Della, & Santandrea, E. (2014). Altering spatial priority maps via reward-based learning. *Journal of Neuroscience*, 34(25), 8594–8604. <https://doi.org/10.1523/JNEUROSCI.0277-14.2014>
- Chelazzi, L., Perlato, A., Santandrea, E., & Della Libera, C. (2013). Rewards teach visual selective attention. *Vision Research*, 85, 58–72. <https://doi.org/10.1016/j.visres.2012.12.005>
- Cheng, F. P. H., Saglam, A., André, S., & Pooresmaeili, A. (2020). Cross-modal integration of reward value during oculomotor planning. *ENeuro*, 7(1). <https://doi.org/10.1523/ENEURO.0381-19.2020>
- Della Libera, C., & Chelazzi, L. (2006). Visual selective attention and the effects of monetary rewards. *Psychological Science*, 17(3), 222–227. <https://doi.org/10.1111/j.1467-9280.2006.01689.x>
- Della Libera, C., & Chelazzi, L. (2009). Learning to attend and to ignore is a matter of gains and losses. *Psychological Science*, 20(6), 778–784. <https://doi.org/10.1111/j.1467-9280.2009.02360.x>
- Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*. <https://doi.org/10.1016/j.jneumeth.2003.10.009>
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, 18, 193–222. <https://doi.org/10.1146/annurev.ne.18.030195.001205>
- Duncan, J., Martens, S., & Ward, R. (1997). Restricted attentional capacity within but not between sensory modalities. *Nature*, 387(6635), 808–810. <https://doi.org/10.1038/42947>
- Failing, M., Nissens, T., Pearson, D., Le Pelley, M., & Theeuwes, J. (2015). Oculomotor capture by stimuli that signal the availability of reward. *Journal of Neurophysiology*, 114(4), 2316–2327. <https://doi.org/10.1152/jn.00441.2015>
- Failing, M., & Theeuwes, J. (2018). Selection history: How reward modulates selectivity of visual attention. *Psychonomic Bulletin & Review*, 25(2), 514–538. <https://doi.org/10.3758/S13423-017-1380-Y/FIGURES/1>
- Garcia-Lazaro, H. G., Bartsch, M. V., Boehler, C. N., Krebs, R. M., Donohue, S. E., Harris, J. A., Schoenfeld, M. A., & Hopf, J. M. (2018). Dissociating reward- and attention-driven biasing of global feature-based selection in human visual cortex. *Journal of Cognitive Neuroscience*, 31(4), 469–481. https://doi.org/10.1162/jocn_a_01356
- Goltstein, P. M., Coffey, E. B. J., Roelfsema, P. R., & Pennartz, C. M. A. (2013). In vivo two-photon Ca²⁺ imaging reveals selective reward effects on stimulus-specific assemblies in mouse visual cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 33(28), 11540–11555. <https://doi.org/10.1523/JNEUROSCI.1341-12.2013>
- 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>
- Hoofs, V., Grahek, I., Boehler, C. N., & Krebs, R. M. (2022). Guiding spatial attention by multimodal reward cues. *Attention, Perception, and Psychophysics*, 84(3), 655–670. <https://doi.org/10.3758/s13414-021-02422-x>
- Hughes, G., Mathan, S., & Yeung, N. (2013). EEG indices of reward motivation and target detectability in a rapid visual detection task. *NeuroImage*, 64(1), 590–600. <https://doi.org/10.1016/j.neuroimage.2012.09.003>
- Kang, G., Wang, L., & Zhou, X. (2017). Reward interacts with modality shift to reduce cross-modal conflict. *Journal of Vision*, 17(1). <https://doi.org/10.1167/17.1.19>
- Kim, A. J., Lee, D. S., & Anderson, B. A. (2021). Previously reward-associated sounds interfere with goal-directed auditory processing. *Quarterly Journal of Experimental Psychology*, 74(7), 1257–1263. <https://doi.org/10.1177/1747021821990033>
- Kiss, M., Driver, J., & Eimer, M. (2009). Reward priority of visual target singletons modulates event-related potential signatures of attentional selection. *Psychological Science*, 20(2), 245–251. <https://doi.org/10.1111/j.1467-9280.2009.02281.x>

- Krebs, R. M., Boehler, C. N., & Woldorff, M. G. (2010). The influence of reward associations on conflict processing in the Stroop task. *Cognition*, 117(3), 341–347. <https://doi.org/10.1016/j.cognition.2010.08.018>
- Le Pelley, M. E., Mitchell, C. J., Beesley, T., George, D. N., & Wills, A. J. (2016). Attention and associative learning in humans: An integrative review. *Psychological Bulletin*, 142(10), 1111–1140. <https://doi.org/10.1037/bul0000064>
- Leo, F., & Noppeney, U. (2014). Conditioned sounds enhance visual processing. *PLoS One*, 9(9), Article e106860. <https://doi.org/10.1371/journal.pone.0106860>
- Luck, S. J., Woodman, G. F., & Vogel, E. K. (2000). Event-related potential studies of attention. *Trends in Cognitive Sciences*, 4(11), 432–440. [https://doi.org/10.1016/S1364-6613\(00\)01545-X](https://doi.org/10.1016/S1364-6613(00)01545-X). Elsevier.
- Lunghi, C., & Pooresmaeili, A. (2023). Learned value modulates the access to visual awareness during continuous flash suppression. *Scientific Reports*, 13(1), 756. <https://doi.org/10.1038/s41598-023-28004-5>
- Maunsell, J. H. R. (2004). Neuronal representations of cognitive state: Reward or attention? *Trends in Cognitive Sciences*, 8(6), 261–265. <https://doi.org/10.1016/j.TICS.2004.04.003>
- Mognon, A., Jovicich, J., Bruzzone, L., & Buiatti, M. (2011). ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features. *Psychophysiology*. <https://doi.org/10.1111/j.1469-8986.2010.01061.x>
- Munneke, J., Belopolsky, A. V., & Theeuwes, J. (2016). Distractors associated with reward break through the focus of attention. *Attention, Perception, and Psychophysics*, 78(7), 2213–2225. <https://doi.org/10.3758/S13414-016-1075-X/FIGURES/4>
- Pessoa, L. (2015). Multiple influences of reward on perception and attention. *Visual Cognition*, 23(1–2), 272–290. <https://doi.org/10.1080/13506285.2014.974729>
- Pleger, B., Blankenburg, F., Ruff, C. C., Driver, J., & Dolan, R. J. (2008). Reward facilitates tactile judgments and modulates hemodynamic responses in human primary somatosensory cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 28(33), 8161–8168. <https://doi.org/10.1523/JNEUROSCI.1093-08.2008>
- Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, 118(10), 2128–2148.
- Pooresmaeili, A., FitzGerald, T. H. B., Bach, D. R., Toelch, U., Ostendorf, F., & Dolan, R. J. (2014). Cross-modal effects of value on perceptual acuity and stimulus encoding. *Proceedings of the National Academy of Sciences of the United States of America*, 111(42), 15244–15249. <https://doi.org/10.1073/pnas.1408873111>
- Qin, N., Gu, R., Xue, J., Chen, C., & Zhang, M. (2021). Reward-driven attention alters perceived salience. *Journal of Vision*, 21(1), 1–13. <https://doi.org/10.1167/JOV.21.1.7>
- Rakhshan, M., Lee, V., Chu, E., Harris, L., Laiks, L., Khorsand, P., & Soltani, A. (2020). Influence of expected reward on temporal order judgment. *Journal of Cognitive Neuroscience*, 32(4), 674–690. https://doi.org/10.1162/jocn_a_01516
- Raymond, J. E., & O'Brien, J. L. (2009). Selective visual attention and motivation: The consequences of value learning in an attentional blink task. *Psychological Science*, 20(8), 981–988. <https://doi.org/10.1111/j.1467-9280.2009.02391.x>
- Rossi, V., Vanlessen, N., Bayer, M., Grass, A., Pourtois, G., & Schacht, A. (2017). Motivational salience modulates early visual cortex responses across task sets. *Journal of Cognitive Neuroscience*, 29(6), 968–979. https://doi.org/10.1162/jocn_a_01093
- Rutkowski, R. G., & Weinberger, N. M. (2005). Encoding of learned importance of sound by magnitude of representational area in primary auditory cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 102(38), 13664–13669. <https://doi.org/10.1073/pnas.0506838102>
- San Martín, R., Appelbaum, L. G., Huettel, S. A., & Woldorff, M. G. (2016). Cortical brain activity reflecting attentional biasing toward reward-predicting cues covaries with economic decision-making performance. *Cerebral Cortex*, 26(1), 1–11. <https://doi.org/10.1093/cercor/bhu160>
- Sanz, L. R. D., Vuilleumier, P., & Bourgeois, A. (2018). Cross-modal integration during value-driven attentional capture. *Neuropsychologia*, 120, 105–112. <https://doi.org/10.1016/j.neuropsychologia.2018.10.014>
- Serences, J. T. (2008). Value-based modulations in human visual cortex. *Neuron*, 60(6), 1169–1181. <https://doi.org/10.1016/j.neuron.2008.10.051>
- Shuler, M. G., & Bear, M. F. (2006). Reward timing in the primary visual cortex. *Science*, 311(5767), 1606–1609. <https://doi.org/10.1126/science.1123513>
- Soltani, A., Rakhshan, M., Schafer, R. J., Burrows, B. E., & Moore, T. (2021). Separable influences of reward on visual processing and choice. *Journal of Cognitive Neuroscience*, 33(2), 248–262. https://doi.org/10.1162/JOCN_A_01647
- Sommer, W., & Schweinberger, S. (1992). Operant conditioning of P300. *Biological Psychology*, 33(1), 37–49. [https://doi.org/10.1016/0301-0511\(92\)90004-E](https://doi.org/10.1016/0301-0511(92)90004-E)
- Stanisor, L., van der Togt, C., Pennartz, C. M. A., & Roelfsema, P. R. (2013). A unified selection signal for attention and reward in primary visual cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 110(22), 9136–9141. <https://doi.org/10.1073/pnas.1300117110>
- Talsma, D., & Woldorff, M. G. (2005). Selective attention and multisensory integration: Multiple phases of effects on the evoked brain activity. *Journal of Cognitive Neuroscience*, 17(7), 1098–1114. <https://doi.org/10.1162/0898929054475172>
- Tankelevitch, L., Spaak, E., Rushworth, M. F. S., & Stokes, M. G. (2020). Previously reward-associated stimuli capture spatial attention in the absence of changes in the corresponding sensory representations as measured with MEG. *Journal of Neuroscience*, 40(26), 5033–5050. <https://doi.org/10.1523/JNEUROSCI.1172-19.2020>
- Vakhrushev, R., Cheng, F. P.-H., Schacht, A., & Pooresmaeili, A. (2023). Differential effects of intra-modal and cross-modal reward value on perception: ERP evidence. *PLoS One*, 18(6), Article e0287900. <https://doi.org/10.1371/journal.pone.0287900>
- Wang, L., Duan, Y., Theeuwes, J., & Zhou, X. (2014). Reward breaks through the inhibitory region around attentional focus. *Journal of Vision*, 14(12). <https://doi.org/10.1167/14.12.2>, 2–2.
- Wang, L., Li, S., Zhou, X., & Theeuwes, J. (2018). Stimuli that signal the availability of reward break into attentional focus. *Vision Research*, 144, 20–28. <https://doi.org/10.1016/j.visres.2017.10.008>
- Weil, R. S., Furl, N., Ruff, C. C., Symmonds, M., Flandin, G., Dolan, R. J., Driver, J., & Rees, G. (2010). Rewarding feedback after correct visual discriminations has both general and specific influences on visual cortex. *Journal of Neurophysiology*, 104(3), 1746–1757. <https://doi.org/10.1152/jn.00870.2009>
- Woldorff, M. G. (1993). Distortion of ERP averages due to overlap from temporally adjacent ERPs: Analysis and correction. *Psychophysiology*, 30(1), 98–119. <https://doi.org/10.1111/j.1469-8986.1993.tb03209.x>
- Wu, Y., & Zhou, X. (2009). The P300 and reward valence, magnitude, and expectancy in outcome evaluation. *Brain Research*, 1286, 114–122. <https://doi.org/10.1016/j.brainres.2009.06.032>
- Yantis, S., Anderson, B. A., Wampler, E. K., & Laurent, P. A. (2012). Reward and attentional control in visual search. In *Nebraska symposium on motivation: Vol. 59. Nebraska symposium on motivation* (pp. 91–116). https://doi.org/10.1007/978-1-4614-4794-8_5. Nebr Symp Motiv.
- Yeung, N., & Sanfey, A. G. (2004). Independent coding of reward magnitude and valence in the human brain. *Journal of Neuroscience*, 24(28), 6258–6264. <https://doi.org/10.1523/JNEUROSCI.4537-03.2004>