

Manuscript Number: NSY-D-18-00286R1

Title: Cross-modal integration during value-driven attentional capture

Article Type: Research Paper

Section/Category: Attention and Consciousness

Keywords: reward; attentional capture; cross-modal integration; manual; saccadic

Corresponding Author: Dr. Leandro R. D. Sanz, M.D.

Corresponding Author's Institution: Coma Science Group

First Author: Leandro R. D. Sanz, M.D.

Order of Authors: Leandro R. D. Sanz, M.D.; Patrik Vuilleumier, Ph.D.; Alexia Bourgeois, Ph.D.

**Abstract:** A growing body of evidence suggests that reward may be a powerful determinant of attentional selection. To date, the study of value-based attentional capture has been mainly focused on the visual sensory modality. It is yet unknown how reward information is communicated and integrated across the different senses in order to resolve between competing choices during selective attention. Our study investigated the interference produced by an auditory reward-associated distractor when a semantically-related visual target was concurrently presented. We measured both manual and saccadic response times towards a target image (drum or trumpet), while an irrelevant sound (congruent or incongruent instrument) was heard. Each sound was previously associated with either a high or a low reward. We found that manual responses were slowed by a high-reward auditory distractor when sound and image were semantically congruent. A similar effect was observed for saccadic responses, but only for participants aware of the past reward contingencies. Auditory events associated with reward value were thus capable of involuntarily capturing attention in the visual modality. This reward effect can mitigate cross-modal semantic integration and appears to be differentially modulated by awareness for saccadic vs. manual responses. Together, our results extend previous work on value-driven attentional biases in perception by showing that these may operate across sensory modalities and override cross-modal integration for semantically-related stimuli. This study sheds new light on the potential implication of brain regions underlying value-driven attention across sensory modalities.



**UNIVERSITÉ  
DE GENÈVE**

**FACULTÉ DE MÉDECINE**

**DÉPARTEMENT DES NEUROSCIENCES FONDAMENTALES**

Campus Biotech - chemin des Mines 9 | Case postale 60 | CH-1211 Genève 20

Tel. +41 (0)22 379 5381 | Fax +41 (0)22 379 5402

Geneva, September 5<sup>th</sup>, 2018

Dear Editor-in-Chief,

- *Manuscript title:* Value-driven attentional capture during cross-modal integration
- *Authors:* Leandro R. D. Sanz, Patrik Vuilleumier, Alexia Bourgeois

Please find enclosed the above-mentioned revised manuscript, which we would like to submit for publication in *Neuropsychologia*.

We have thoroughly read the detailed comments from the two reviewers whom we thank for their careful reading and their insightful suggestions. We provide here a revised version of our manuscript addressing all the issues and keypoints raised by the reviewers.

Thanking you in advance for your attention,  
With best regards,

On behalf of all co-authors  
Leandro R. D. Sanz

A handwritten signature in black ink, appearing to read 'L. Sanz'.

Ms. No.: NSY-D-18-00286

Title: Cross-modal integration during value-driven attentional capture

Corresponding Author: Dr. Leandro R. D. Sanz

Authors: Patrik Vuilleumier, Ph.D.; Alexia Bourgeois, Ph.D.

## REVIEWER 1

The paper by Sanz and colleagues describes a research aimed at investigating the behavioral correlates of the effects of reward on multi-sensory integration. In particular, the authors measured the interference produced by irrelevant, lateralized auditory stimuli (which were previously associated with a high vs. low reward during a dedicated learning phase) on the ability of the participants to discriminate the spatial location of a subsequent, lateralized visual target. Critically, in different experimental conditions, auditory and visual stimuli were either semantically congruent or incongruent (semantic congruency), and were presented either on the same or different spatial hemifield (spatial congruency). During the test phase, participants identified the lateralization of the visual target with both a manual and an oculomotor response. An increased interference by auditory distractors was assessed on manual responses in semantically congruent conditions; on the contrary, no effect was observed on oculomotor responses, except for participants who had become aware of stimulus-reward contingencies during the training phase.

The topic is of interest and the paper is generally well written and pleasant to read. However, a number of disputable methodological choices and a suboptimal, and sometimes misleading, approach to data analyses undermine the overall quality of the work and its potential impact. In what follows, I will point to some major weaknesses that the authors should try to address in a revised version of the manuscript. Some minor points will instead be listed at the end of this review, together with some indications of missing references that should be covered.

*Response: We thank the reviewer for their careful reading and kind positive suggestions. Below we addressed the different concerns expressed by the reviewer and changes made in the manuscript.*

## MAIN ISSUES: DESIGN

The paradigm described by the authors has some non-trivial points of weakness.

First, the authors aim at measuring attentional capture effects on both manual and oculomotor responses during the test phase. Their choice was to require both responses on the very same trial, probably for the valuable reason that in this way they were able to test both effectors on the very same subjects (although no within-subjects analysis was then applied to directly make comparisons). However, this choice is highly disputable for different reasons:

*Response: The reviewer expresses concerns about the choice of testing both effectors on the same trial and in a within-subject design. We agree that this decision may require some cautions regarding the interpretation of the results obtained on each effector (i.e. manual and saccadic). However, this choice was precisely motivated to investigate the interdependency of reward-related effect on both manual and saccadic responses, which would have been impossible to achieve in a single-effector paradigm. Given the potential individual variability in reward integration, a within-subject design was chosen to compare more reliably the influence of reward on each effector.*

*We understand that the rationale behind the paradigm design as well as its limitations were unclear in the initial version of the manuscript and the text has been modified to address these issues.*

a) It is not clear whether correct responses for both effectors are required to obtain a reward at the end of each trial - based on what is written in the methods, I think only the manual response was considered

to determine accuracy in performance, which makes the two responses clearly different in terms of their relevance to task execution.

***Response:** The reviewer is right, the accuracy in performance was only based on manual responses. Note however that the feedback given to participants during the testing phase was to ensure that participants correctly performed the task throughout the experiment and was not related anymore to the reward manipulation. Participants were clearly informed on the flat reward system in the testing phase and instructed to only try and respond as fast as possible using both effectors. We acknowledge that the absence of feedback for the oculomotor response may have induced an imbalance in task relevance between the two effectors. However, we believe that a double feedback (manual and oculomotor) would have induced confusion and may have complicated the realization of the task. The percentage of correct manual responses was very high (98.6%). Only 54 trials corresponding to an incorrect manual response (out of 4370 trials included in the analysis) and a consecutive correct saccade have been discarded across all participants, suggesting that the concomitant oculomotor task may not impact manual accuracy and may not impact the pattern of results observed in this study.*

***Action:** A table with the accuracy for each effector has been added in the revised version of the manuscript (see Supplementary Material, new Table S1). The text has been modified to state more clearly that only trials with correct manual and saccadic responses were considered for the analysis (see revised manuscript, p. 15 l. 11), and a paragraph in the Discussion has been added to stress this potential limitation (see revised manuscript, p. 22 l. 10-16)*

b) The oculomotor and manual responses are not coordinated (e.g., are not aimed to the same target location), which makes them potentially conflicting in terms of motor planning.

***Response:** Manual and oculomotor responses always had to be performed towards the image, i.e. towards the same target location.*

***Action:** We apologize for the lack of clarity regarding this point. The Methods section has been reformulated to clarify (see revised manuscript, p. 10 l. 4-8).*

c) While manual responses were employed in the association (training) phase in relation to the discrimination of sound lateralization, eye movements were required during the test phase anew. It would have been much better to test the effects of reward associations on manual and oculomotor responses in separate group of subjects/experiments, with both used as the effective response to the task. Repeating the experiment in this form will be important to make any claim about differential patterns for manual vs. saccadic responses tenable.

***Response:** The absence of oculomotor response during the association phase might have indeed limited reward-induced effects for this effector. However, our main purpose was to evaluate how semantically-related cross-modal stimuli associated with a reward could modulate manual responses. The oculomotor effector was used as a secondary endpoint to investigate a possible interaction between motor and oculomotor systems during value-driven attentional capture. Additionally, a within-subject design was considered more suitable regarding the high individual variability in reward integration.*

***Action:** A paragraph has been added in the Discussion to deal with this point (see revised manuscript, p. 22 l. 16-18).*

Second, the choice of using a flat, low-reward schedule during the test phase is highly disputable too. Although in the latter phase fixed rewards are associated with responses to visual targets (not to sounds), a strong reduction in reward availability, within a very similar motivational context, might have had an unspecific impact on the participant attitude. Since unbalanced rewards were associated with specific sounds in the initial training phase, one could hypothesize that a more general association was created between specific objects (semantic level) and reward levels. If this were the case, the test phase, with flat

low reward availability will have served as a strong extinction protocol, negating the previously learnt associations, or at least promoting unbinding of multisensory information. Extinction will be in place even if one only considers that a purely associative (Pavlovian) link was established between sounds and rewards. Although the authors should have better used a no-reward protocol during the test phase, now they should at least try to see whether reward effects might be unveiled more clearly during the initial blocks of the test phase, where perhaps the effects of extinction will not be complete yet.

*Response: As stated above (see response a), the feedback given to participants during the testing phase was to ensure that participants correctly performed the task and was not related anymore to reward manipulation. Participants were clearly instructed in this sense before each phase. We apologize for this potential lack of clarity. Nevertheless, we agree that the similar motivational context of the testing phase compared to the association phase could have elicited some extinction phenomenon. We could however assume that extinction would have abolished rather than elicited reward-related effects during the testing phase, as suggested by our results.*

*In order to deal with this point, we have conducted the analyses on separate blocks of trials to search for progressive extinction.*

*Action: This additional analysis has been added in the Results section (see revised manuscript, p. 14 l. 10-17) and discussed in the section 4 (see revised manuscript, p. 21 l. 6-10). We have only found a main effect of trial blocks such that participants were faster during the last trials (376ms) compared to the first ones (430ms), as can be expected with training. No interaction between trial blocks and any other factor was significant.*

Finally, based on how the task is designed, participants are required to make use of attentional resources to avoid/override multisensory integration (ignore irrelevant stimuli in the auditory modality and select relevant stimuli in the visual modality). The authors should make explicit mention this aspect and to the expected influence of an association of rewards which is confined to a selectively auditory dimension or instead might have acquired a broader semantic meaning.

*Response: The reviewer is right. The task was indeed designed to investigate how previously rewarded stimuli modulated attentional performances in the visual modality when target and distractor were semantically related. We agree with the reviewer that two possible reward learning patterns may have been expected, depending on whether the value was purely associated with the auditory stimulus or rather with the semantic concept of the instrument.*

*Action: We now have explicitly mentioned the distracting nature of the auditory stimulus in the Methods section (see revised manuscript, p. 9 l. 9 and p. 10 l. 4) and discussed the possible interpretations both in the Introduction (see revised manuscript, p. 6 l. 18-19) and in the Discussion section (see revised manuscript, p. 20 l. 3-9).*

## **MAIN ISSUES: ANALYSIS**

There are also some relevant problems in the way the authors approach data analyses. While the approach to manual RTs is generally sound, the approach to saccadic RTs appears as incomplete and more confused, perhaps also because of the complexity retrieved in the actual pattern of results (see below).

Before coming to actual results, a general issue pertains to the approach to saccadic responses. Did the authors consider only first saccades in their analyses?

Response: We indeed considered in the analysis the first saccade falling in a ROI defined as a circle of 6° of visual angle around the target picture's center. This methodology allowed us to focus on goal-directed saccades without considering saccades not directed towards the target, and hence reliably computing saccadic response times of eye movements directed towards the target.

Action: The Section 3.2.2 has been modified to clarify this point (see revised manuscript, p. 15 l. 8-10).

Why did the authors consider only saccadic responses in trials where a correct manual response was collected?

Response: As stated above, the percentage of correct manual responses was very high (98.6%). Only 53 trials corresponding to an incorrect manual response (out of 4370 trials included in the analysis) and a consecutive correct saccade have been discarded across all participants, suggesting that the concomitant oculomotor task may not impact manual accuracy and may not impact the pattern of results observed in this study. Also, manual responses were the main endpoint in the design of this study, using oculomotor responses as a secondary measure.

Action: This point has been clarified in the revised manuscript (see revised manuscript, p.22 l. 14-16).

Were there trials in which a "correct" saccade (first saccade to the target) was measured in the absence of a correct manual response? The authors should first report general data describing the percentage of first saccades correctly landing on the visual target, saccadic errors or omissions and the behavioral relations between the two kinds of responses they measured during the task.

Response: We thank the reviewer for this suggestion.

Action: We have now added a table reporting percentages of accuracy for manual and saccadic responses (see new Table S1, p. 29). It is important to mention that very few trials (53 trials across all participants) contained an incorrect manual response times and a valid saccade. These exclusions might thus not impact the pattern of results observed in this study.

Coming to results collected with saccadic responses, the authors describe an overall effect of awareness, an interaction between awareness and reward, and a quadruple interaction between all factors. They then decide to concentrate on a rather specific comparison, namely they measure the effect of reward on the attentional capture by a semantically-congruent and location-incongruent sound distractor in aware vs. unaware participants. This is interesting in principle, but the authors should better justify their focus on this aspect and, to provide a suitable parallel, they should also show results of a similar analysis for manual responses.

Response: We thank the reviewer for this comment. The rationale behind this analysis was that the condition where the sound and the image from the same instrument presented in a different hemispace corresponded to the highest possible interference. This situation was therefore probed for an effect of awareness that may be hidden in less interferential conditions. We agree that the clarity of the methodology would benefit from a similar analysis on manual response times.

Action: We have now performed the same analysis on manual response times. No significant effect was observed for this effector (see revised manuscript, p. 17 l. 16-22).

Moreover, I think that other potentially interesting effects are disregarded. Just as an example, for both aware and unaware participants, an interaction between spatial and semantic congruency for high reward stimuli appears to be evident from inspection of Figure 4. Was this significant? I suggest that the authors

provide a more articulated account of the significant results observed with saccadic RTs (decomposing all potentially meaningful aspects of the quadruple interaction).

Response: *We agree with the reviewer that this complex quadruple interaction needed more detailed account in the manuscript.*

Action: *We have decomposed the full analysis and performed four ANOVAs separating aware from unaware participants, in high reward and in low reward conditions. We found a significant interaction between semantic congruence and location congruence for highly rewarded trials in participants aware of the reward association, as suggested by the reviewer,  $F(1,7)=8.49$ ,  $MSE=73$ ,  $p=0.023$ . No other main effect or interaction in any of the four ANOVAs was significant (all  $ps > 0.07$ ). We have reported these additional analyses in the text (see revised manuscript, p. 17 l. 1-6).*

Finally, and most critically, the authors should report correlations between the critical effects found with manual responses and saccadic responses (also in the case of non-significant effects in one or the other case) in order to understand whether the effects found in the two domains do share common substrates or have an independent origin/nature.

Response: *We agree with the reviewer that correlations between manual and saccadic responses can bring further information on the interdependence between effectors and help understand the differential results observed.*

Action: *We have performed Pearson correlations between manual and saccadic RTs in the testing phase for similar experimental conditions, which showed no significant correlation for any of the critical conditions showing a significant main effect or interaction in one or the other modality (all  $ps > 0.06$ ). These additional analyses have been reported in the text (see revised manuscript, p. 18 l. 1-4).*

In addition, in a related perspective, effects measured during the test phase should be correlated with the results of the association phase.

Response: *We fully agree with the reviewer.*

Action: *We have performed Pearson correlations between the manual RTs in the association and the testing phase. After the exclusion of one participant who showed exceptionally slow RTs in the association phase, this analysis revealed a significant positive correlation between manual RTs in the association and the testing phase ( $R=0.69$ ,  $p=0.001$ ). Additional correlations considering only trials with a highly-rewarded sound ( $R=0.64$ ,  $p=0.001$ ) or a low-rewarded sound ( $R=0.72$ ,  $p=0.003$ ) also revealed significant positive correlations between the association and the testing phase. These results confirm the consistency of subjects' performance across the two phases of this experiment. These additional analyses have been reported in the text (see revised manuscript, p. 14 l. 18-22 and p. 15 l. 1-2).*

## MINOR POINTS

### INTRODUCTION:

Additional references are needed when addressing the influence of top-down signals on different sensory processing stages (e.g., Buschman & Kastner, Neuron 2015; Chelazzi et al., WIRE 2011, among others). Additional reference will also be appropriate in relation to the description of reward effects as a form of attentional control (e.g., Failing and Theeuwes, Psychon Bull Rev 2018; Chelazzi et al., Vis Res 2013; Awh et al., Trends Cog Sci 2012 as reviews, again among others).

Response: *We thank the reviewer for these suggestions.*

Action: *These references have been added in the revised manuscript (see p. 3 l. 17, p. 5 l. 12, p. 4 l. 18).*

In the last part of the Introduction (p.5), the authors cite opposite findings in the literature (Pooresmaeili et al., 2014 and Anderson, 2015b); it would be important to underline and explain the contradictions also in view of a deeper discussion of the authors' own findings, as reported in the current paper.

Response: We agree with the reviewer. Note that the experimental design used in these two studies contained important differences, which could explain the differential effect of reward associated stimuli observed.

Action: *The text has been modified to further discuss the results of these two studies (see revised manuscript, p. 6 l. 3-8)*

## METHODS:

Please clarify how the high vs. low reward outcome was justified to the participants (as dependent from their performance? as delivered randomly?).

Response: *We apologize for the lack of clarity regarding this point. Participants were instructed to maximize the number of points that they could win throughout the experiment, without being informed of the rule underlying reward contingencies and probabilities.*

Action: *This statement has been clarified in the revised manuscript (see revised manuscript, p. 8 l. 12-14).*

Please report the refresh rate of the monitor (section 2.2, p. 7).

Response: We acknowledge the omission of this information.

Action: *We have reported the monitor refresh rate in the revised manuscript (see p. 7 l. 17).*

The temporal distance between the sound and the visual stimulus during the test phase was fixed at 200 ms, rendering the sound a very precise temporal cue. In addition, the sound might have served as a spatial cue to automatically bias attention, which in half of the trials (spatially incongruent trials) would likely have produced a cost in performance. Indeed, the inter-stimulus distance was such to avoid inhibition of return. All these aspects pertaining to task design should be made explicit and discussed.

Response: *The reviewer is right, a constant 200ms stimulus onset asynchrony (SOA) was chosen in order to avoid inhibition of return effect. Only the location congruence of the sound and the image could vary. Our results indicated slower responses, although marginally significant, when the sound and the image were spatially incongruent.*

Action: *This aspect has been mentioned and clarified in the revised manuscript (see p. 9 l. 13-14 and p. 10 l. 1-2).*

Please clarify if reward in the test phase was delivered only on the basis of manual responses (see also above).

Response: *As indicated above, no reward manipulation was performed during the testing phase. Instead, feedback regarding the accuracy of manual responses was given during this phase, to ensure that participants performed the discrimination task throughout the experiment correctly.*

Action: *This statement has been clarified in the revised manuscript (see p. 10 l. 10-11).*

## RESULTS:

A figure is needed for readers to visualize the overall significant interaction between block and reward in the association phase (p. 9-10, section 3.1).

*Response: We thank the reviewer for this suggestion.*

*Action: A figure has been added in the revised manuscript (see new Figure S1, p.29).*

For manual responses, reported analyses should be anyway repeated without "aware" participants to fully confirm the reported pattern of results.

*Response: We thank the review for the suggestion and agree with the necessity to confirm results with this analysis.*

*Action: We have repeated the ANOVA on manual RTs including only participants unaware of the association with reward. Similarly as observed with all participants, an interaction between reward and semantic congruence was found,  $F(1,11)=6.20$ ,  $MSE=289$ ,  $p=0.030$ . No other main effect or interaction was present. A mention to this analysis was added to the text (see p. 14 l. 7-9).*

## DISCUSSION:

To discuss the peculiarities in the pattern of results observed for eye movements, it will be interesting to focus on what reported in the main part of this review (lack of task relevance for the saccadic response).

*Response/Action: The Section 4 has been modified in order to take into account the lack of task relevance for the saccadic task, which could have explained the pattern of results observed in this study (see revised manuscript, p. 22 l. 10-23).*

The idea that reward effects might be overall concealed on saccadic responses because the paradigm did not involve reflexive saccades is not that convincing. Manual responses, which were modulated by reward signals, were not automatic movements and would have likely suffered from a similar problem. Also, the fact that an effect on saccades was found in aware participants is not enough to sustain that perspective; perhaps this aspect might push to an account of the results based on considering how strategy might have affected saccadic responses in aware participants (once again related to the previous issue).

*Response: The Discussion part has been modified to integrate this suggestion (see revised manuscript, p. 23 l. 15-23).*

In relation to the main topic of the paper, i.e. how cross-modal integration is affected by reward-mediate attentional capture, the authors should consider and discuss recent papers which might be of great relevance for their own research (e.g., Bruns et al., *Atten Percept Psychophys* 2014; Talsma et al., *Trends Cog Sci* 2010; Macaluso, *Cortex* 2010).

*Response: We thank the reviewer for these relevant additional references, which have been integrated in the revised manuscript (see p. 4 l. 1&14, p. 5 l. 23, p. 6 l. 18, p. 18 l. 15, p. 20 l. 3).*

## REVIEWER 2

In this study Sanz and colleagues aim to explore whether and how the effects of value-based attentional capture can be observed across different sensory modalities. What I find particularly interesting about this approach is that, differently from previous studies in the literature, the auditory and visual stimuli that have been employed here are coupled from a semantic point of view. Therefore this study actually investigates whether and how reward associations with specific stimuli delivered in the auditory affect the processing of the same objects, perceived through a visual image.

*Response: We thank the reviewer for their positive comments and suggestions. Below we addressed the different concerns expressed by the reviewer and changes made in the manuscript.*

This difference relatively to previous studies on stimulus-reward associations in the auditory or visual domain should be stressed more clearly in the introduction. At present, especially in the last part of the

introduction, when previous studies on reward and auditory stimuli are discussed (e.g., Pooriesmaeili et al., 2014 and Anderson, 2015b), it is somewhat difficult to grasp the important novel aspects which are addressed in this study.

*Response: We acknowledge that the multisensory integration and semantic coupling of stimuli were not particularly highlighted in our introduction despite their novelty in our design.*

*Action: The introduction has been modified in order to more clearly underline novel aspects addressed in this study and its potential impact on our understanding of reward-related effect on attentional perception (see revised manuscript, p. 6 l. 15-19, p. 7 l. 1-2, p. 19 l. 2).*

Here indeed the different stimulus modalities refer intrinsically to the same concept, therefore whether or not crossover effects will be found due to reward-associations depends on whether reward will affect "central" stimulus representations, rather than simple, low-level auditory or visual signals.

*Response: The reviewer is right. The task was indeed designed to investigate how previously rewarded stimuli modulated attentional performances in the visual modality when target and distractor were semantically related. We agree with the reviewer that two possible reward learning patterns may have been expected, depending on whether the value was purely associated with the auditory stimulus or rather with the semantic concept of the instrument.*

*Action: We now have explicitly mentioned this aspect and discussed the possible interpretations both in the Introduction (see revised manuscript, p. 6 l. 18-19) and in the Discussion section (see revised manuscript, p. 20 l.3-9).*

To this respect I found that it is crucial also to stress that the tasks that subjects are required to perform NEVER entail stimulus discrimination, either in the learning or in the test phase. Access to stimulus identity is never needed, as responses should be delivered according to stimulus location, and yet, whether the effects will appear will depend on an association between reward and stimulus identity.

*Response: The reviewer is right. The task was indeed designed to avoid access to stimulus identity that could have potentially interfered with the cross-modal semantic integration of visual and auditory stimuli.*

*Action: The Discussion has been modified to highlight this important aspect (see revised manuscript, p. 19 l. 13-15).*

## **Results:**

As far as I could expect a priori, the main result of the study is somewhat contradictory. If the processing of reward-associated stimuli is enhanced, and the level of processing at which reward acts is semantic, one might have expected that trials with congruent stimuli in the test should have received a better performance, and especially so for those trials in which the spatial location of the two is also congruent (with respect to when locations are opposite). Instead, not only there is no impact of past reward-associations on spatial congruency, but the effect that emerges is quite the opposite. Irrespectively of stimulus locations, trials in which the rewarded signal is presented together with the congruent visual stimulus lead to a worsening of performance, as if the rewarded signal were able to trigger competition, rather than facilitation, with respect to the processing of the semantically associated visual item.

*Response: The reviewer is right. The semantic identity shared by auditory distractors and visual targets allows to test if the value was purely associated with the auditory stimulus or rather with the semantic concept of instrument. An integration of the reward value at a semantic level would have been likely to*

*elicit enhanced performance in semantically congruent trials with highly rewarded sounds. The opposite pattern of results observed here suggests that value is rather associated with the auditory stimulus itself, which induces an attentional competition with semantic binding processes.*

Action: *Discussion on this consideration has been added to the revised version of the manuscript (see p.20 l. 3-9).*

This finding was not corroborated by saccadic latencies towards the visual stimuli, which were only affected by semantic congruency altogether, being faster to the onset of visual stimuli that were semantically congruent to the auditory ones. In my view this effect was probably driven by the fact that auditory stimuli preceded in time visual onsets, and therefore might have created a semantic context which affected attentional deployment. This possibility, never acknowledged, should be taken into consideration in the Discussion.

Response: *We fully agree with the reviewer.*

Action: *This hypothesis is now mentioned in the Discussion (see revised manuscript, p. 22 l. 18-23)*

In addition to these findings, the Authors also report that performance across subjects was rather influenced by participants' awareness of stimulus-rewards associations during the learning session. Indeed, in saccadic latencies subjects that had become aware of reward contingencies ( $n = 8$  individuals) showed significant costs in performance in trials in which the visual and auditory stimuli were semantically congruent, but spatially incongruent. This finding is actually the only one in line with what one could have expected on the basis of prior literature on value-based attentional capture.

I wonder whether part of the mismatch between the results obtained from manual RTs and saccadic latencies could be explained by the fact that subjects were explicitly required to provide both responses. Such requirement might have given rise to an attempt to control more strategically the timing of the two responses (e.g., eye-movement first, key-press later), and this extra effort could perhaps have overshadowed the effects that might have appeared (in manual responses) in a more "natural" context, i.e., when saccades were spontaneous and not required.

Instead of taking this possibility into consideration, the Authors discuss in length the possibly dissociable neural mechanisms underlying eye- and key-press responses. While I think that this might actually be one of the possible causes for this dissociation, one way to corroborate this hypothesis could be to prove that they are NOT correlated to one another. Are they? These tests should be made, and the findings obtained discussed.

Response: *We fully agree that the timing of the two responses may have influenced the pattern of results observed in this study. As suggested by the reviewer, we calculated statistical correlations between manual and saccadic response times. This analysis indicated no significant correlation (all  $ps > 0.06$ ), suggesting the existence of distinct neural mechanisms underlying manual and saccadic reward interference.*

Action: *The Results and Discussion parts have been modified to discuss this point (see revised manuscript, p. 18 l. 1-4 and p. 22 l. 21-23).*

Overall the manuscript is well written and easy to follow, but in general, I find that much part of the introduction and of the discussion deal with describing broadly neural networks and mechanisms that cannot be really pinpointed with the current results.

Given the nature of the task and of the findings obtained, I suggest that more emphasis should be put on speculating about the cognitive processes that are tackled by these manipulations, rather than the neural

underpinnings. As the authors suggest, these might be addressed in a future (perhaps neuroimaging) study.

*Response: We recognize that the results presented in the present study allow only speculative conclusions on the neural bases of cross-modal integration during value-driven attentional capture. In order to meet the reviewer concern, we have modified section 1 and 4 to reduce and mitigate this speculative assumptions (see revised manuscript, p.24 l. 8-11, p. 7 l. 2, p. 20 l. 11, p. 20 l. 14 & l. 19, deleted paragraph at p. 19 l. 5).*

### **Figure 1**

The task is well described, however some of the elements in the picture are not discernible. For the purpose of illustration, fixation crosses and arrows indicating the trial timeline should be rendered thicker.

*Response: We thank the reviewer for this comment. These modifications have been made (see new Figure 1, p. 9).*

### **Graphs**

Overall I found the graphs to be somewhat confusing, as they plot the raw data even when the full blown interaction is non-significant. Perhaps it would be a good idea to plot means at the levels that give rise to significant effects, so that the message carried by the statistical analysis does not get diluted in the particular non-significant trends shown at higher levels of detail. Table 1 already provides the means for all conditions and sublevels, so all one needs to know is already there.

*Response: We agree with the reviewer's suggestion. We have removed the details regarding location congruence in all graphs to give a better understanding of significant effects.*

*We would like to thank the reviewer for their insightful suggestions and comments about our study.*

# Cross-modal integration during value-driven attentional capture

Leandro R. D. Sanz<sup>a,b</sup>, Patrik Vuilleumier<sup>a\*</sup>, Alexia Bourgeois<sup>a\*</sup>

*\* Both authors equally contributed to this manuscript*

- <sup>a.</sup> Neuroscience Department, Laboratory for Behavioral Neurology and Imaging of Cognition, University of Geneva, Geneva, Switzerland.
- <sup>b.</sup> Coma Science Group, GIGA Consciousness, University and University Hospital of Liège, Liège, Belgium

## Highlights

- How is reward information integrated across the different sensory modalities?
- Rewarded auditory distractors capture attention during visual target processing.
- Reward interferes with manual responses if image and sound share semantic identity.
- Manual responses are interfered by reward signals without conscious awareness.
- Conscious awareness modulates cross-modal value-driven effects on saccadic responses.

1                   **Cross-modal integration during value-driven attentional**  
2  
3  
4                   **capture**  
5  
6

7  
8                   3                   Leandro R. D. Sanz<sup>a,b 1</sup>, Patrik Vuilleumier<sup>a \*</sup>, Alexia Bourgeois<sup>a \*</sup>  
9

10                   4  
11  
12                   5  
13  
14                   6                   \* *Both authors equally contributed to this manuscript*  
15

- 16  
17                   7                   a.   Neuroscience Department, Laboratory for Behavioral Neurology and  
18                   8                   Imaging of Cognition, University of Geneva, Geneva, Switzerland.  
19                   9  
20                   10                   b.   Coma Science Group, GIGA Consciousness, University and University  
21                   11                   Hospital of Liège, Liège, Belgium  
22  
23  
24  
25  
26

27                   14                   Patrik Vuilleumier, patrik.vuilleumier@unige.ch  
28

29                   15                   Alexia Bourgeois, alx.bourgeois@gmail.com  
30  
31  
32                   16

33  
34                   17                   Corresponding author: Leandro R. D. Sanz, leandro.sanz@uliege.be, +32 489 33 84 91  
35  
36                   18  
37  
38                   19  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58

59                   1                   <sup>1</sup> Present address : Coma Science Group, GIGA Consciousness, University and University Hospital of Liège,  
60                   Avenue de l’Hôpital 1, Liège, Belgium  
61

## Abstract

A growing body of evidence suggests that reward may be a powerful determinant of attentional selection. To date, the study of value-based attentional capture has been mainly focused on the visual sensory modality. It is yet unknown how reward information is communicated and integrated across the different senses in order to resolve between competing choices during selective attention. Our study investigated the interference produced by an auditory reward-associated distractor when a semantically-related visual target was concurrently presented. We measured both manual and saccadic response times towards a target image (drum or trumpet), while an irrelevant sound (congruent or incongruent instrument) was heard. Each sound was previously associated with either a high or a low reward. We found that manual responses were slowed by a high-reward auditory distractor when sound and image were semantically congruent. A similar effect was observed for saccadic responses, but only for participants aware of the past reward contingencies. Auditory events associated with reward value were thus capable of involuntarily capturing attention in the visual modality. This reward effect can mitigate cross-modal semantic integration and appears to be differentially modulated by awareness for saccadic vs. manual responses. Together, our results extend previous work on value-driven attentional biases in perception by showing that these may operate across sensory modalities and override cross-modal integration for semantically-related stimuli. This study sheds new light on the potential implication of brain regions underlying value-driven attention across sensory modalities.

Keywords: reward, attentional capture, cross-modal integration, manual, saccadic

Declarations of interest: none

## 1. Introduction

Despite the highly efficient and parallel processing capacity of the human brain, our ability to encode and respond to external sensory information is limited. Selective attention provides an adaptive filtering system that allows for selecting relevant stimuli and ignoring irrelevant information in the environment. However, attention is not a unitary mechanism. Unexpected, salient stimuli can capture attention in an involuntary, bottom-up manner through exogenous orienting mechanisms. Or instead, attention can be oriented in space in a top-down, goal-directed manner, through endogenous control mechanisms (Posner, 1980). These two components of attention rely on distinct neuroanatomical networks: while endogenous attention is primarily mediated by activation in the intraparietal and superior frontal cortices, exogenous orienting activates more ventral areas in temporoparietal and inferior frontal cortices. These two systems entertain highly interactive functional relationships, with the latter acting as a circuit breaker on the former when salient stimuli require reorienting and refocusing of attention (Corbetta and Shulman, 2002). Neuroscience research has consistently shown that these frontoparietal networks mediate attention selection mechanisms by imposing top-down modulatory signals on sensory pathways (Buschman and Kastner, 2015; Chelazzi et al., 2011; Corbetta and Shulman, 2002; Quentin et al., 2015), leading to enhanced neuronal responses at several processing stages from primary up to higher association stages (Vuilleumier et al., 2001).

Information coming from different sensory modalities also competes for attentional capacity and thus implies cross-modal integration mechanisms (Jolicoeur, 1999). While multisensory integration has often been considered as an automatic process, some evidence suggested that it could

1 be modulated by attentional processing (Talsma et al., 2010). Well-known attentional  
2 effects observed in unimodal conditions are also encountered across different sensory  
3 inputs. For instance, a visual target shortly following an auditory target generates an  
4 attentional blink of a similar size as observed with unimodal visual targets. (Arnell and  
5 Larson, 2002). Spatial cueing benefits described in the visual modality are also elicited by  
6 cross-modal cues presented in the auditory modality prior to a visual target (McDonald et  
7 al., 2012; Spence, 2010 for reviews). Furthermore, conflicting information from different  
8 modalities can generate attentional biases which can modify perception, as shown by the  
9 ventriloquist illusion where synchronous visual and auditory information are presented in  
10 close spatial proximity, giving the impression that the sound comes from the visual stimulus  
11 (Vroomen and De Gelder, 2004 for a review). Neuroimaging evidence suggests that cross-  
12 modal integration relies on an increase of activity in corresponding unimodal sensory  
13 cortices, mediated by top-down signals from multimodal areas (Hillyard et al., 2016; Johnson  
14 and Zatorre, 2005; Macaluso, 2010, 2000; Shomstein and Yantis, 2004), in ways partly similar  
15 to attentional effects mediated by frontoparietal networks (Driver and Noesselt, 2008).

16 Interestingly, recent literature suggests that attentional selection is also strongly  
17 modulated by value-associated signals that code for the motivational or affective  
18 significance of stimuli (Anderson, 2015a; Chelazzi et al., 2013; Vuilleumier, 2015). A visual  
19 feature previously associated with a high reward can subsequently automatically capture  
20 attention in a task where this feature is irrelevant (Anderson et al., 2013, 2011a; Bourgeois et  
21 al., 2017; Della Libera and Chelazzi, 2006). This effect may operate independently of intrinsic  
22 stimulus saliency, task goals, or strategic set (Anderson et al., 2011b; Failing and Theeuwes,  
23 2014). These value-driven mechanisms resemble both selective top-down and stimulus-  
24 driven attentional orienting, as they depend on internal states but arise without conscious

1 will (Bourgeois et al., 2017; Hickey et al., 2014, 2010; Vuilleumier and Brosch, 2009). They  
2 seem however functionally and anatomically distinct from attentional systems classically  
3 associated with frontoparietal cortical networks (Anderson et al., 2014; Lucas et al., 2013;  
4 Serences, 2008). Thus, there is evidence that visual reward information is communicated  
5 through a ‘value-driven attention network’ encompassing early visual cortical areas  
6 (Macleod and Giesbrecht, 2015; van Koningsbruggen et al., 2016), extrastriate visual cortex  
7 (Anderson et al., 2014), the intraparietal sulcus (Peck et al., 2009; Qi et al., 2013), as well as  
8 the tail of the caudate nucleus (Hikosaka et al., 2014) and the substantia nigra pars reticulata  
9 (Yasuda and Hikosaka, 2015). These signals appear to be further encoded in areas implicated  
10 in saccade generation such as the superior colliculus (Ikeda and Hikosaka, 2007) and frontal  
11 eye field (FEF; Ding and Hikosaka, 2006), which could allow the modulation of spatial priority  
12 maps and generate reward-based biases in gaze control (Awh et al., 2012; Failing and  
13 Theeuwes, 2018).

14       Reward may elicit attentional capture by acting on different sensory pathways and  
15 effector systems. While most studies used manual responses to track attentional effects of  
16 reward-associated distractors, eye movements are also powerfully affected by the reward  
17 value of visual stimuli, with changes in both the latency and trajectory of saccades during  
18 search or detection tasks (Hickey and van Zoest, 2013; Theeuwes and Belopolsky, 2012). This  
19 is consistent with reward signals recorded in superior colliculus and frontoparietal areas  
20 controlling oculomotor behavior (Hikosaka et al., 2014; Ikeda and Hikosaka, 2007). Further,  
21 while most research on value-based attentional capture focused on the visual modality,  
22 cross-modal effects have been described by few recent studies. For instance, ventriloquist  
23 effect has been found to be reduced in the hemifield associated with a high reward  
24 expectancy compared to the low-reward hemifield (Bruns et al., 2014). Highly-rewarded

1 task-irrelevant auditory stimuli could enhance perceptual sensitivity in a visual  
2 discrimination task when presented simultaneously with targets (Pooresmaeili et al., 2014).  
3 This effect was correlated with more accurate representation of the target stimulus  
4 orientation in primary visual cortex, as well as with cross-modal activations within the  
5 temporal lobe. In a different task based on the detection of a shape singleton in a search  
6 array, an auditory distractor previously associated with a high reward could interfere with  
7 the detection of a paired visual target (Anderson, 2015b). These contrasting results  
8 underline that reward can modulate different components of visual attention through  
9 facilitation or interference, depending on the task set and the type of stimuli used.

10       How reward information is communicated and integrated across different sensory  
11 modalities in order to resolve between competing inputs remains, however, poorly  
12 understood. To date, the postulated neural correlates of value-driven attention have been  
13 mostly confined to the visual system, due to the predominance of studies relying on the  
14 visual modality (Anderson et al., 2014; Hickey and Peelen, 2015). Demonstrating reward  
15 interference using auditory stimuli would provide direct evidence that reward learning  
16 broadly impacts attentional mechanisms beyond the visual system. Moreover, no study has  
17 investigated how reward-driven biases modulate the integration of multisensory information  
18 sharing semantic identity. Notably, while reward learning effects during multisensory  
19 integration have been described (Bruns et al., 2014), it is unknown whether value is encoded  
20 at a primary sensory level or rather at higher integrative semantic integrative level.  
21 Additionally, cross-modal reward effects driven from auditory cues on eye movements have  
22 never been documented.

Here we therefore recorded both manual and saccadic response times (RTs) in a visual orienting task to investigate the interference produced by reward-associated auditory distractors on the detection of a visual target which could share semantic features with the sound. Our results will shed new light on the cognitive processes underlying value-driven attention across sensory modalities and the integration of reward signals in the brain.

## **2. Methods**

### **2.1 Participants**

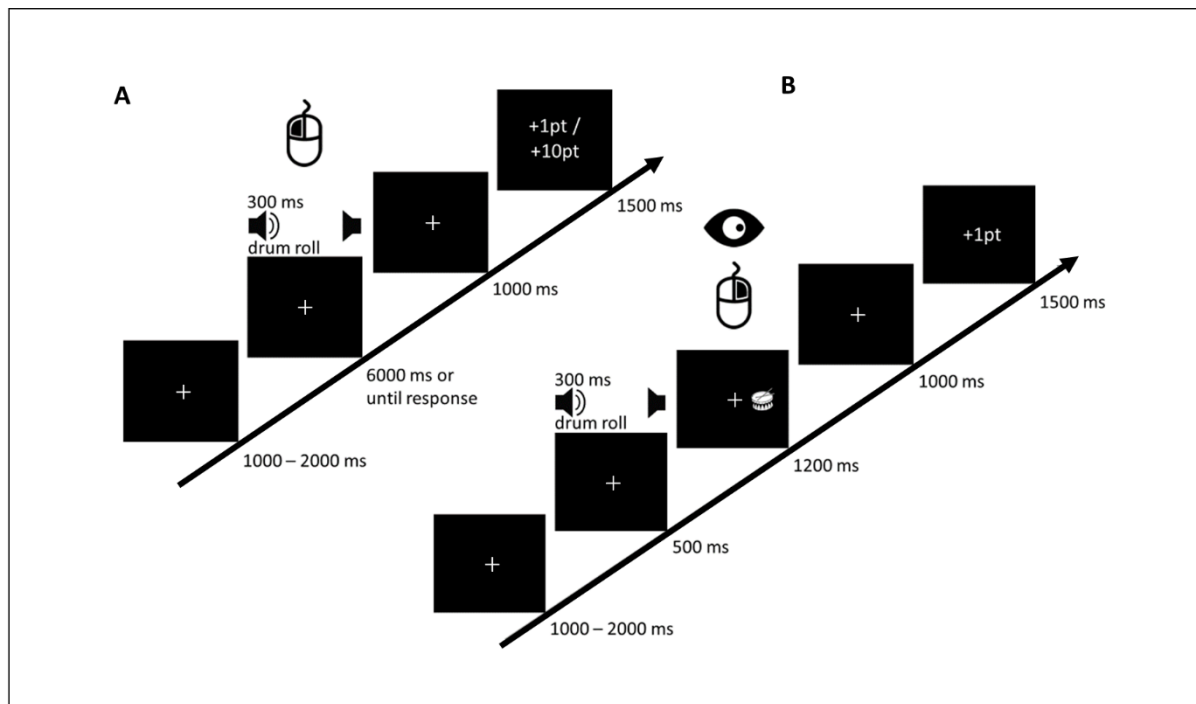
Twenty healthy volunteers (eight women, all right-handed, mean age 24.8 years, range 19-28) with normal or corrected-to-normal vision and audition, as well as no history of neurological or psychiatric disorders, participated in this study. Written informed consent was obtained from each participant, according to procedures approved by the local ethical committee. They were compensated with a 10 CHF flat rate.

### **2.2 Apparatus, stimuli and procedure**

A PC running E-prime 2.0 software (Psychology Software Tools, Sharpsburg, USA; Schneider et al., 2002) controlled the presentation of stimuli, timing operations, and data collection. An Eyelink 1000 device (SR Research, Ottawa, Canada) was used to record monocular left eye position with a 1000 Hz resolution. Participants rested their head on a chin-rest 57 cm from the monitor (DELL P2210, Round Rock, USA; screen resolution 1280 x 1024, refresh rate 60Hz). Auditory stimuli, with sound level equalized across participants, were presented via headphones (Sennheiser HD380, Wedemark, Germany).

The task consisted in a learning / association phase followed by a testing phase. Both phases comprised 240 trials and were introduced to our participants as two unrelated

1 experiments on perception. During the initial learning / association phase, participants had  
2 to report the lateralization of sounds and received reward feedback after their response. On  
3 each trial, a central white fixation cross ( $0.8^\circ \times 0.8^\circ$  of visual angle) was first presented  
4 against a black background for a pseudo-random interval ranging from 1000 to 2000 ms,  
5 followed by a unilateral 300 ms sound of either a trumpet or a drum roll (equalized for  
6 intensity). Participants were asked to discriminate the lateralization of the sound with a  
7 corresponding button click (left or right click), as fast and as accurately as possible. After  
8 their response, or after 6000 ms if no response was recorded, a visual feedback informed  
9 participants about the monetary reward earned on that trial, together with the total reward  
10 accumulated across all preceding trials. One of the two instruments (trumpet or drum roll,  
11 counterbalanced across participants) was followed by a high reward feedback (10 points,  
12 presented together with a picture of golden coins) on 80% of correct trials, and a low reward  
13 feedback (one point) on the remaining 20%. These percentages were reversed for the low-  
14 rewarded instrument. Incorrect responses were given 0 points. Participants were instructed  
15 to maximize the number of points that they could win throughout the experiment, without  
16 being informed of the rule underlying reward contingencies and probabilities (Figure 1A).



**Figure 1.** Experimental task set. Sequence of events and timing of a trial during the learning / association phase (A) and subsequent testing phase (B). A mouse symbol is pictured when a manual response is required and an eye symbol is pictured when a saccadic response is required.

After a short break, participants performed the testing phase. This was similar to the initial association phase, with the following changes. After the fixation cross, a lateralized 300 ms sound of trumpet or drum roll was presented. This auditory distractor was followed, 200 ms later, by the presentation of a black-and-white image depicting a drum or a trumpet (3° of visual angle), displayed on either the right or the left of the fixation cross, in the center of the hemiscreen. This delay between sound and image allowed for the sound identity to unfold completely and thus effectively compete for attentional selection. Moreover, a constant 200ms stimulus onset asynchrony (SOA) was chosen in order to avoid inhibition of

1 return effect. Therefore, the sound could automatically bias attention and produce a cost in  
2 performance in spatially incongruent trials. Participants were now required to discriminate  
3 as fast and as accurately as possible the lateralization of the image with a corresponding  
4 button click, while ignoring the auditory distractor. To further assess the attentional capture  
5 by a previously high-rewarded sound for both manual and eye movement responses,  
6 participants were also required to perform a saccade towards the same target picture, as  
7 soon as it appeared, in addition to the manual button click. Both manual and oculomotor  
8 responses were thus performed towards the same target location. In order to ensure that  
9 the task was correctly performed, correct manual responses were given one point, while  
10 incorrect responses were given 0 points. No feedback was given on the accuracy of the  
11 oculomotor response. Participants were instructed to respond as fast and as accurately as  
12 possible for both manual and saccadic responses. (Figure 1B).

13 Participant's debriefing reports were collected at the end of the experiment to probe  
14 whether they had guessed the association between rewards and auditory targets in the first  
15 learning / association phase.

### 16 **3. Results**

#### 17 **3.1 Association phase**

18 In this phase, participants had to judge the lateralization of an auditory stimulus  
19 representing either a drum or a trumpet, with only one of these two instruments associated  
20 with a high reward (counterbalanced across participants). However, the sound identity was  
21 irrelevant to the task. We analyzed correct responses with RTs inside a three standard  
22 deviations (SD) interval (95.6% of total trials). A repeated-measure analysis of variance  
23 (ANOVA) on mean RTs using the within-participant factor of reward (high vs. low) showed no

significant difference for the highly rewarded sound (473 ms) compared to the low rewarded sound (472 ms),  $F(1,19)=0.022$ ,  $MSE=456$ ,  $p=0.883$ .

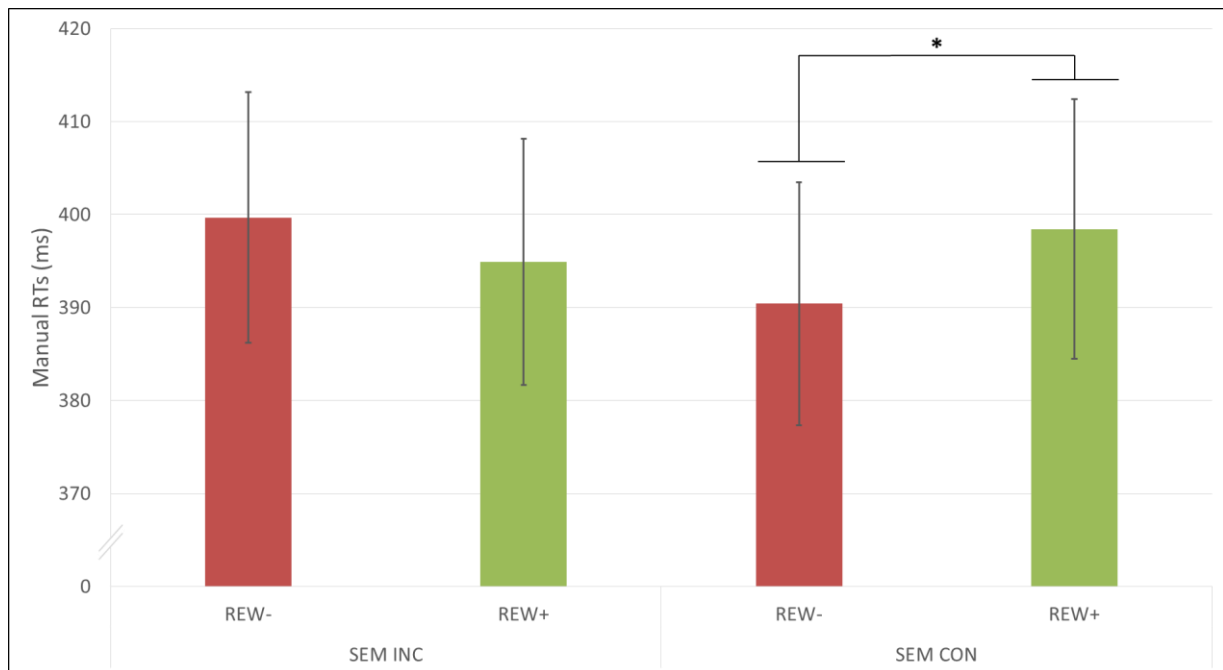
To examine the effect of reward learning over time, we divided these data into five blocks of 48 consecutive trials each and performed a new ANOVA with the within-participant factors of reward (low, high) and blocks (1 to 5). This analysis showed no main effect of blocks,  $F(4,76)=1.75$ ,  $MSE=4239$ ,  $p=0.154$ , but a significant interaction between reward and blocks,  $F(4,76)=2.51$ ,  $MSE=2034$ ,  $p=0.049$ . While RTs were very similar in the first block (low reward, 448 ms; high reward, 455 ms), we observed numerically faster RTs to respond to high-reward compared to low-reward targets for the following blocks of trials, except for the last block where this pattern of RTs reversed (Figure S1). However, direct pairwise comparisons indicated no significant differences between high and low reward-associated stimuli for any block (all  $ps > 0.12$ ).

### 3.2 Testing phase

In this second phase, participants had now to discriminate the lateralization of an image depicting either a drum or a trumpet, presented briefly after a lateralized sound of either a drum or a trumpet. Half of trials were drum pictures, and the other half trumpet pictures. Orthogonally to this, the sound and image were semantically congruent in 50% of trials (e.g., drum picture presented after the sound of a drum roll) and incongruent in 50% (e.g., drum picture after trumpet sound). Thus, in half of the trials, the sound was previously associated with a high reward (during the first association phase), but with a low reward in the remaining half. In addition, the image target appeared in the same hemispace as the sound (location congruent) or on a different hemispace (location incongruent) in half of trials each.

### 3.2.1 Manual RTs

We recorded manual RTs from the onset of the visual target until a button click was made. Only correct responses with RTs inside a three SD interval were considered for analysis (96.5% of total trials). We then submitted mean RTs to a repeated-measure analysis of variance (ANOVA) with the within-participant factors of reward (previously high vs. low rewarded sound), semantic congruence (sound and image of same vs. different instruments), and location congruence (sound and image in the same vs. opposite hemispace). This analysis revealed a significant interaction between reward and semantic congruence,  $F(1,19)=5.90$ ,  $MSE=281$ ,  $p=0.025$ . Participants showed slower RTs when the sound was previously associated with a high reward, specifically for trials with semantic congruence ( $p=0.047$ ), but there was no such difference for incongruent trials ( $p=0.214$ ) (Figure 2). Thus, the interference effect of reward on RTs was present only when sound and image shared semantic identity. Also, participants showed a trend to respond faster when the sound and image were presented in the same hemispace rather than in a different hemispace,  $F(1,19)=3.40$ ,  $MSE=1169$ ,  $p=0.081$ , consistent with classic spatial compatibility effects. No other main effect or interaction was significant (all  $ps > 0.18$ ; Figure 2 and Table 1).



**Figure 2.** Manual responses. Mean manual RTs (ms) obtained during the testing phase as a function of whether visual targets were semantically incongruent (SEM INC) or congruent (SEM CON) with auditory distractors and whether auditory stimuli were previously associated with a low (REW-) or high (REW+) reward. \*  $p < 0.05$

	SEM INC				SEM CON			
	REW- LAT INC	LAT CON	REW+ LAT INC	LAT CON	REW- LAT INC	LAT CON	REW+ LAT INC	LAT CON
Manual responses	404 (15)	395 (12)	400 (16)	389 (12)	394 (13)	387 (14)	404 (15)	393 (13)
Saccadic responses	335 (8)	340 (7)	335 (6)	334 (7)	328 (6)	330 (7)	328 (6)	331 (8)

**Table 1.** Mean manual and saccadic RTs as well as standard deviations indicated in parentheses (in ms) for visual targets semantically incongruent (SEM INC) or congruent (SEM CON) with auditory distractors, for auditory stimuli previously associated with a low (REW-) or high (REW+) reward, and for visual targets presented in a different (LAT INC) or the same hemispace (LAT CON).

1 In final debriefing, eight out of 20 participants reported they had correctly guessed the  
2 association between reward and sound in the association task. In order to investigate the  
3 effect of awareness on subsequent reward effects, we repeated our ANOVA on the mean  
4 manual RTs as above but now adding awareness (aware vs. unaware) as a between-  
5 participants factor. Importantly, this analysis did not indicate any significant interaction  
6 between awareness and experimental conditions (all  $ps > 0.27$ ), suggesting that reward  
7 effects did not depend on conscious control for manual responses. In addition, identical  
8 results were obtained when repeating the ANOVA on the mean manual RTs without patients  
9 aware of the association between reward and sound, which corroborates this hypothesis.

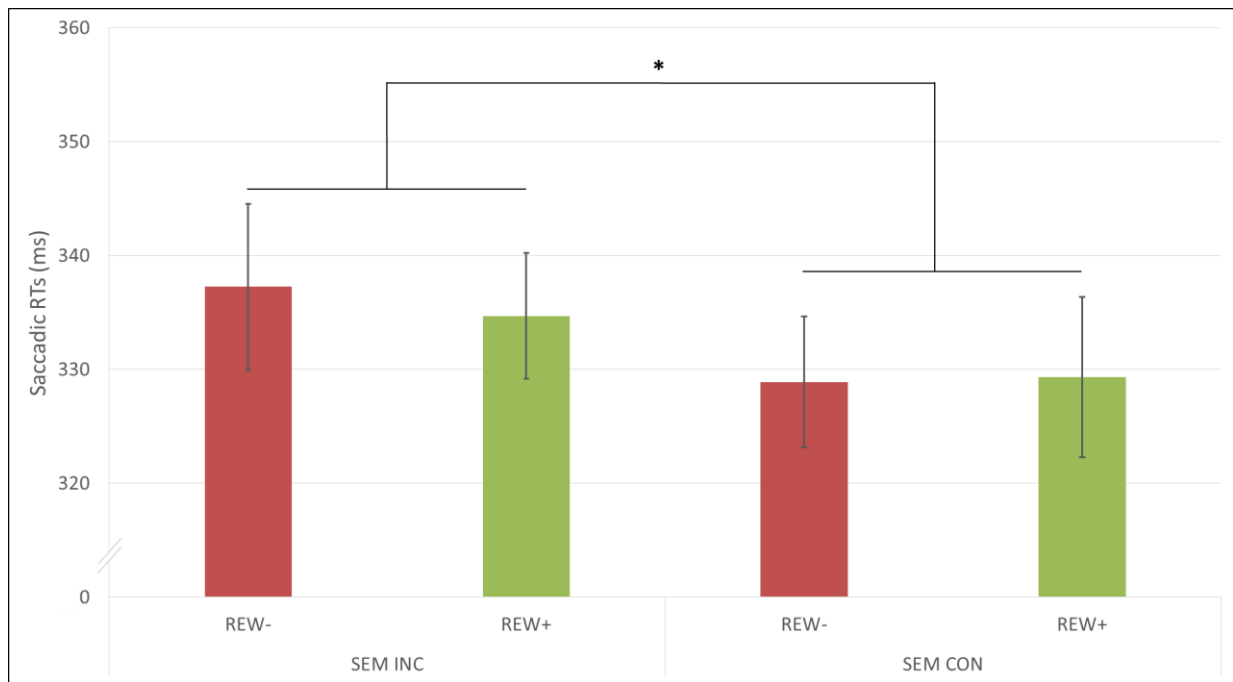
10 To probe the effect of learning over time in the testing phase, we again divided the  
11 trials in five blocks of 48 trials, as described above for the association phase, and performed  
12 a new ANOVA adding the within-participant factor of trial blocks. We found a significant  
13 main effect of trials blocks,  $F(4,76)=9.87$ ,  $MSE = 7214$ ,  $p < 0.001$ . Participants were faster  
14 during the last trials (376 ms) compared to the first ones (430ms), as can be expected with  
15 training. Planned comparisons between the first and the last block of trials confirmed these  
16 results ( $p<0.001$ ). No other significant main effect or interaction between trial blocks and  
17 other experimental conditions was found, all  $ps > 0.12$ .

18 Finally, the interdependence of subjects' performance across phases was probed using  
19 Pearson correlation coefficients between manual RTs during the association phase and  
20 during the testing phase. After the exclusion of one outlier subject with slower RTs in the  
21 association phase, a significant positive correlation was found between manual RTs in the  
22 association and the testing phase ( $R=0.69$ ,  $p=0.001$ ). Additional analyses considering only  
23 trials with a highly-rewarded sound ( $R=0.64$ ,  $p=0.001$ ) or a low-rewarded sound ( $R=0.72$ ,

1 p=0.003) also revealed significant positive correlations between the association and the  
2 testing phase.

### 3.2.2 Saccadic RTs

Concomitantly to their manual responses to report the image location, participants were instructed to make a saccade towards the target picture. For each trial, saccadic RT was recorded from the target image onset till the end of the first saccade falling inside a region of interest (ROI) overlapping with the target picture. This ROI was defined as a circle of 6° of visual angle around the picture's center. Considering the first saccade ending inside the ROI, which was not necessarily the first saccade of the trial, allowed us to target only goal-directed saccades and reliably compute saccadic response times. Only trials with correct manual response, at least one saccade ending in the correct ROI and a saccadic RTs inside a 3 SD interval were included in the analysis (90.5% of total trials). We then performed an ANOVA on mean saccadic RTs with the within-participant factors of reward (previously high vs. low rewarded sound), semantic congruence (sound and image of same vs. different instruments), and location congruence (sound and image presented in the same vs. opposite hemispace). This showed again a main effect of semantic congruence,  $F(1,19)=5.02$ ,  $MSE=372$ ,  $p=0.037$ , with faster saccades when the image and the sound were from the same instrument. However, unlike for manual RTs, this effect was not modulated by reward ( $p=0.702$ ). No other main effect or interaction was significant (all  $ps > 0.41$ ; Figure 3 and Table 1).



**Figure 3.** Saccadic responses. Mean saccadic RTs (ms) obtained during the testing phase as a function of whether visual targets were semantically incongruent (SEM INC) or congruent (SEM CON) with auditory distractors and whether auditory stimuli were previously associated with a low (REW-) or high (REW+) reward. \*  $p < 0.05$

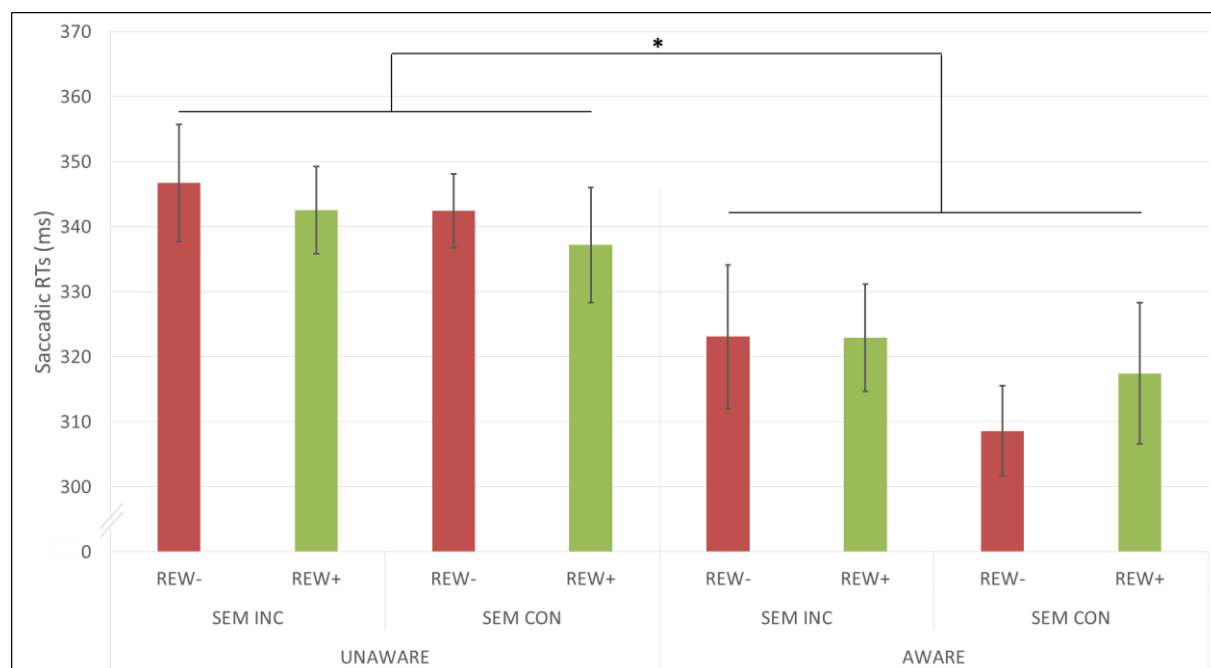
As previously done for the manual RTs analysis, we also ran an ANOVA with the additional between-participant factor of awareness (i.e., participants who noticed the preferential association of reward with one sound in the first phase). This analysis revealed a significant main effect of awareness,  $F(1,18)=5.01$ ,  $MSE=4486$ ,  $p=0.038$ : saccades were generally faster in aware than unaware participants. More critically, the interaction between awareness and reward was also significant,  $F(1,18)=7.09$ ,  $MSE=124$ ,  $p=0.016$ , as well as the quadruple interaction between reward, awareness, semantic congruence, and location congruence,  $F(1,18)=6.50$ ,  $MSE=151$ ,  $p=0.020$  (Figure 4).

1 To better understand this quadruple interaction, we decomposed the analysis and  
2 performed four separate ANOVAs separating aware from unaware participants, in high  
3 reward and in low reward conditions. We found a significant interaction between semantic  
4 congruence and location congruence for highly rewarded trials in participants aware of the  
5 reward association,  $F(1,7)=8.49$ ,  $MSE=73$ ,  $p=0.023$ . No other main effect or interaction in any  
6 of the four ANOVAs was significant (all  $ps > 0.07$ ).

7 To help the interpretation of these results, we calculated an index of attentional  
8 capture by previously rewarded sounds, as follows: saccadic RTs in presence of the high-  
9 reward sound minus saccadic RTs in presence of the low-reward-sound, in the conditions  
10 with the strongest potential interference, i.e. trials where distractors were semantically  
11 congruent but location incongruent. We then performed a two-tailed  $t$ -test against 0 to  
12 define any significant interference or facilitation by the reward-associated sound. Results  
13 indicated that participants who were aware of the association exhibited a significant  
14 interference cost due to the high-reward sound ( $p=0.05$ , index mean=13.3ms), while  
15 participants unaware of the association showed no significant effect ( $p=0.24$ , index mean=-  
16 9.5ms).

17 For completeness, an attentional capture index was calculated using the same method  
18 for manual responses in the testing phase. We first calculated the index on all participants  
19 regardless of awareness, in conditions of spatial congruency and location incongruency, and  
20 then repeated similar analyses considering separately subjects aware and unaware of  
21 reward associations. We performed a two-tailed  $t$ -test of this index against 0. Neither the  
22 whole population ( $p=0.07$ ) nor any of the subgroups (aware,  $p=0.29$ ; unaware,  $p=0.16$ )  
23 showed an attentional capture index significantly different from zero in these conditions.

To corroborate the independence of results between the two effector modalities, Pearson correlations were calculated between manual and saccadic RTs in the testing phase for all experimental conditions where a significant main effect or interaction was identified. No significant correlation was found for any of these experimental conditions (all  $p > 0.06$ ).



**Figure 4.** Awareness and saccadic responses. Mean saccadic RTs (ms) obtained during the testing phase as a function of whether participants were unaware or aware of reward contingencies, whether visual targets were semantically incongruent (SEM INC) or congruent (SEM CON) with auditory distractors, whether auditory stimuli were previously associated with a low (REW-) or high (REW+) reward, and whether visual targets were presented in a different (LAT INC) or the same hemispace (LAT CON) as auditory distractors. \*  $p < 0.05$

#### 4. Discussion

Although value-driven biases in attention have been shown to act across different sensory modalities (Anderson, 2015b; Bruns et al., 2014; Pooresmaeili et al., 2014), the influence of reward on cross-modal integration mechanisms remains largely unresolved, in

1 particular when multisensory stimuli share semantic identity. Moreover, to the best of our  
2 knowledge, whether these mechanisms differentially impact on orienting responses  
3 measured through hand or eye movements has not been investigated. Here we directly  
4 tested how reward history can bias attention to stimuli sharing common semantic features,  
5 using visual targets that could either match or not match with the category of a co-occurring  
6 sound associated with different reward values, while we recorded both manual and saccadic  
7 RTs.

8       When participants were required to report the location of a visual image with manual  
9 responses, our results demonstrated that previously rewarded distractors interfered with  
10 cross-modal semantic binding. Indeed, highly rewarded distractors elicited delayed manual  
11 responses only when they shared semantic identity with the target image. This indicates that  
12 the high reward value of the auditory distractor could hinder its cross-modal semantic  
13 integration with the visual stimulus, leading to distraction by the sound and slower RTs. This  
14 effect was independent of stimulus identification since participants were never required to  
15 discriminate the target during both phases of the experiment.

16       Such slowing by high-reward sounds accords with previous results showing that value-  
17 associated stimuli can capture attention regardless of their task-relevance or intrinsic  
18 sensory salience (Anderson et al., 2011b; Bourgeois et al., 2017; Della Libera and Chelazzi,  
19 2006; Failing and Theeuwes, 2014). Most of the past research, however, used visual stimuli  
20 as both targets and distractors (e.g. in visual search tasks), in which different visual features  
21 are paired with different reward values through learning (Bourgeois et al., 2016a for a  
22 review). Our novel data further extend these findings by demonstrating that, first, value-  
23 driven biases in attention can also operate through auditory stimuli, and second, they can

1 influence the processing of stimuli during cross-modal integration, which corroborates the  
2 inference that multisensory integration can be influenced by attentional effects induced by  
3 reward-associated stimuli (Bruns et al., 2014; Talsma et al., 2010). Moreover, the semantic  
4 identity shared by auditory distractors and visual targets allowed to test if the value was  
5 purely associated with the auditory stimulus or rather with the semantic concept of  
6 instrument. An integration of the reward value at a semantic level would have been likely to  
7 elicit enhanced performance in semantically congruent trials with highly rewarded sounds.  
8 The opposite pattern of results observed here suggests that value is rather associated with  
9 the auditory stimulus itself, which induced an attentional competition with semantic binding  
10 processes. These results may provide new insights on the neural bases underlying value-  
11 driven attention across sensory modalities. Indeed, value-based modulation has been mainly  
12 observed for visual stimuli within early sensory regions of the visual cortex (Anderson et al.,  
13 2014; Serences, 2008). Our results provide strong evidence that reward learning exerts a  
14 broad impact on attention that extends beyond the visual system. Moreover, as reward  
15 signals seem to induce an interference that can override semantic cross-modal association,  
16 we surmise that they may act through rapid and automatic mechanisms modulating early  
17 sensory processing, prior to higher stages of semantic integration. In line with reward-  
18 related effects observed in the visual domain (Hickey and Peelen, 2015), we speculate that  
19 auditory reward signals might also impact sensory processing within the auditory cortex. This  
20 would be consistent with previous research showing that value-driven effects in attention  
21 are usually effective even without awareness of differential reward associations (Bourgeois  
22 et al., 2016b), as we found for manual RTs here too. Future research is needed to confirm  
23 these assumptions and further investigate the neural bases of reward-associated stimuli in  
24 the auditory domain.

1 More generally, our findings add to other recent studies where multimodal stimuli  
2 were also used, though with important task differences. For instance, previously rewarded  
3 spoken letters produced an attentional capture in a visual search array task with letters as  
4 targets and distractors (Anderson, 2015b). Analogous interference effects were observed in  
5 our study but for targets and distractors presented in different modalities, and only when  
6 they were semantically-congruent. Unlike these previous results, our task did not show a  
7 main effect of reward on performance. This discrepancy may be due either to our more  
8 complex task set which featured semantic binding of stimuli, either to possible extinction  
9 phenomena during the testing phase due to the flat-reward point system used. The latter is  
10 however not corroborated by the analysis performed on trial blocks.

11 Moreover, our results demonstrated interference effects of reward-associated stimuli,  
12 in contrast with a previous study where the high value of task-irrelevant sounds could  
13 enhance the perceptual sensitivity of visual targets (Pooresmaeili et al., 2014). These  
14 contrasting results may be explained by the absence of semantic relationship between the  
15 visual targets (Gabor stimuli) and the rewarded sound (pure beeps) in this latter study.

16 In our study, participants also made a saccadic response toward visual targets.  
17 Remarkably, an effect of semantic congruence between auditory and visual stimuli on  
18 performance was also observed. However, this bias did not seem to be modulated by  
19 reward. Production of saccades is known to be mainly under the control of the frontal eye  
20 field for intentional saccades and the posterior eye field for reflexive saccades, with both  
21 areas sending inputs to the superior colliculi that integrates and encodes saccade  
22 coordinates (Pierrot-Deseilligny et al., 2003). Each of these structures at the cortical and  
23 subcortical levels have been shown to be directly modulated by reward information (Ding

and Hikosaka, 2006; Ikeda and Hikosaka, 2007). However, unlike studies testing reward effects with unimodal visual stimuli, our task involved cross-modal integration between reward-associated auditory stimuli and non-rewarded visual targets. This additional step may imply a delay that could prevent value-driven biases to act on purely reflexive mechanisms underlying saccades, and thus could explain the absence of a general reward-induced effect on saccadic responses in our paradigm. Furthermore, we found no systematic slowing of saccadic responses when a previously rewarded sound was presented on the side opposite to the visual targets, suggesting that value-based biases in attention orienting in our task were perhaps not driven by direct reward signals to the superior colliculus and the basal ganglia circuit (i.e. the caudate nucleus and the substantia nigra; Yasuda and Hikosaka, 2015). This differential pattern of results might also be explained by a lack of relevance of the saccadic task compared to the manual one. Indeed, only feedback regarding the accuracy of the manual responses was provided to participants, which might have increased the focus on this response modality. However, the high accuracy for both manual (98.6%) and saccadic (92.1%) responses suggests that the concomitant tasks did not impact performance and thus the observed pattern of results (Table S1). Also, only manual responses were required during the association phase. This could have potentially facilitated reward-related effects on this specific effector and hindered reward-related effects on oculomotor responses. Moreover, the facilitation of saccadic RTs by semantic congruence may also be explained by the sequential timing of the two stimuli that could have induced a semantic context enhancing attentional deployment after the auditory stimulus. Nevertheless, the absence of statistical correlation between the two effectors rather suggests that the observed results derive from the existence of distinct neural mechanisms underlying manual and saccadic reward interference.

On the other hand, we found that participants who were aware of the previous reward contingencies exhibited slower saccades in presence of high-rewarded sound compared to a low-rewarded sound (i.e., value-driven attentional capture), specifically on semantically congruent trials, whereas unaware participants did not show this effect. These data suggest that awareness could facilitate the appraisal of reward information during saccade programming, leading to stronger interference by the distractor sound and thus slower saccades. These data corroborate previous results that also observed stronger attentional biases by reward in an attentional orienting task for subjects aware of reward contingencies (Bourgeois et al., 2017), even though reward-induced effects can also frequently operate without conscious awareness (Bourgeois et al., 2016a). These patterns of performance might reflect the establishment of an endogenous attentional set with greater top-down control among aware participants, amplifying reward-induced biases. As manual responses did not show such effect of awareness, we surmise that this differential impact on saccadic responses might reflect a functional overlap in anatomical structures such as FEF for endogenous attention and eye movement control. Our results might also at least partly be explained by differences regarding strategic priorities between manual and saccadic responses. As stated above, value-driven biases for saccadic responses may have been less prioritized compared to manual responses. However, this strategic task set could have been modulated by conditions of awareness, leading to the observation of reward-related effect on saccadic responses in patients aware of saccadic contingencies selectively. Alternatively, our results might also be explained by an attempt to control more strategically the timing of the two responses, which could have elicited an extra attentional cost and overshadowed reward-based effects for saccadic responses.

1 Altogether, our results demonstrate an attentional capture elicited by auditory  
2 distractors previously associated with reward, which can modulate concomitant cross-modal  
3 integration effects during visual target detection. Value learning interferes with both manual  
4 responses and eye movements when visual targets shared semantic identity with rewarded  
5 auditory distractors. However, interference on saccadic responses occurred only when  
6 participants were aware of reward contingencies, suggesting a strengthening of value-driven  
7 biases on spatial orienting and eye movements under conscious awareness conditions.

8 Future research may investigate the neural correlates of value-driven attentional  
9 capture across different sensory modalities using neuroimaging techniques such as  
10 functional magnetic resonance imaging, in order to further understand how reward signals  
11 are integrated and communicated in the brain.

## 12 **Acknowledgments**

13 This work was supported by grants from the Fondation Fyssen (AB); the Marie-Curie  
14 CoFund BRIDGE program (AB) from the European Union Seventh Framework Programme  
15 (FP7/2007-2013) under grant agreement n° 267171; from the Swiss National Science  
16 Foundation (SNF no. 162744 and no.166704 to PV), the National Center of Competence in  
17 Research (NCCR) for the Affective Sciences financed by the Swiss National Science  
18 Foundation (SNF No 51NF40-104897) and hosted by the University of Geneva; an award  
19 from the Geneva Academic Society (Foremane Fund to PV); and a grant from the Human  
20 Brain Project (EU-H2020-FETFLAGSHIP-HBP-SGA1-GA720270) to LS.

## 21 **References**

- 22 Anderson, B.A., 2015a. The attention habit: how reward learning shapes attentional  
23 selection. *Ann. N. Y. Acad. Sci.* <https://doi.org/10.1111/nyas.12957>  
24 Anderson, B.A., 2015b. Value-driven attentional capture in the auditory domain. *Attention,*

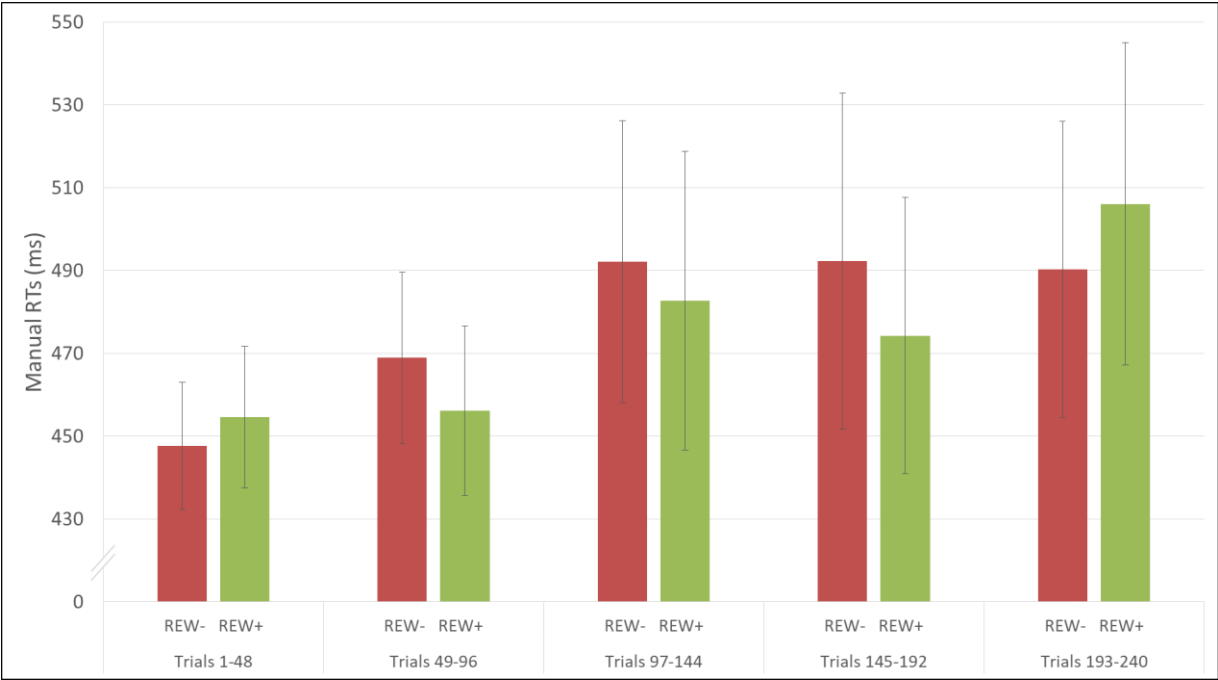
- 1 Perception, Psychophys. <https://doi.org/10.3758/s13414-015-1001-7>
- 2 2 Anderson, B.A., Laurent, P.A., Yantis, S., 2014. Value-driven attentional priority signals in
- 3 3 human basal ganglia and visual cortex. *Brain Res.* 1587, 88–96.
- 4 4 <https://doi.org/10.1016/j.brainres.2014.08.062>
- 5 5 Anderson, B.A., Laurent, P.A., Yantis, S., 2013. Reward predictions bias attentional selection.
- 6 6 *Front Hum Neurosci* 7, 262. <https://doi.org/10.3389/fnhum.2013.00262>
- 7 7 Anderson, B.A., Laurent, P.A., Yantis, S., 2011a. Learned Value Magnifies Salience-Based
- 8 8 Attentional Capture. *PLoS One* 6, e27926.
- 9 9 <https://doi.org/10.1371/journal.pone.0027926>
- 10 10 Anderson, B.A., Laurent, P.A., Yantis, S., 2011b. Value-driven attentional capture. *Proc. Natl.*
- 11 11 *Acad. Sci.* 108, 10367–10371. <https://doi.org/10.1073/pnas.1104047108>
- 12 12 Arnell, K.M., Larson, J.M., 2002. Cross-modality attentional blinks without preparatory task-
- 13 13 set switching. *Psychon. Bull. Rev.* 9, 497–506. <https://doi.org/10.3758/BF03196305>
- 14 14 Awh, E., Belopolsky, A. V., Theeuwes, J., 2012. Top-down versus bottom-up attentional
- 15 15 control: A failed theoretical dichotomy. *Trends Cogn. Sci.* 16, 437–443.
- 16 16 <https://doi.org/10.1016/j.tics.2012.06.010>
- 17 17 Bourgeois, A., Chelazzi, L., Vuilleumier, P., 2016a. How motivation and reward learning
- 18 18 modulate selective attention. *Prog. Brain Res.* 229, 325–342.
- 19 19 <https://doi.org/10.1016/bs.pbr.2016.06.004>
- 20 20 Bourgeois, A., Neveu, R., Bayle, D.J., Vuilleumier, P., 2017. How does reward compete with
- 21 21 goal-directed and stimulus-driven shifts of attention? *Cogn. Emot.* 31, 109–118.
- 22 22 <https://doi.org/10.1080/02699931.2015.1085366>
- 23 23 Bourgeois, A., Neveu, R., Vuilleumier, P., 2016b. How does awareness modulate goal-
- 24 24 directed and stimulus-driven shifts of attention triggered by value learning? *PLoS One*
- 25 25 11, 1–13. <https://doi.org/10.1371/journal.pone.0160469>
- 26 26 Bruns, P., Maiworm, M., Röder, B., 2014. Reward expectation influences audiovisual spatial
- 27 27 integration. *Attention, Perception, Psychophys.* 76, 1815–1827.
- 28 28 <https://doi.org/10.3758/s13414-014-0699-y>
- 29 29 Buschman, T.J., Kastner, S., 2015. From Behavior to Neural Dynamics: An Integrated Theory
- 30 30 of Attention. *Neuron* 88, 127–144. <https://doi.org/10.1016/j.neuron.2015.09.017>
- 31 31 Chelazzi, L., Della Libera, C., Sani, I., Santandrea, E., 2011. Neural basis of visual selective
- 32 32 attention. *Wiley Interdiscip. Rev. Cogn. Sci.* 2, 392–407.
- 33 33 <https://doi.org/10.1002/wcs.117>
- 34 34 Chelazzi, L., Perlato, A., Santandrea, E., Della Libera, C., 2013. Rewards teach visual selective
- 35 35 attention. *Vision Res.* 85, 58–62. <https://doi.org/10.1016/j.visres.2012.12.005>
- 36 36 Corbetta, M., Shulman, G.L., 2002. Control of goal-directed and stimulus-driven attention in
- 37 37 the brain. *Nat. Rev. Neurosci.* 3, 201–215. <https://doi.org/10.1038/nrn755>

- 1 Della Libera, C., Chelazzi, L., 2006. Visual selective attention and the effects of monetary  
2 rewards. *Psychol. Sci.* 17, 222–227. <https://doi.org/10.1111/j.1467-9280.2006.01689.x>
- 3 Ding, L., Hikosaka, O., 2006. Behavioral/Systems/Cognitive Comparison of Reward  
4 Modulation in the Frontal Eye Field and Caudate of the Macaque. *J. Neurosci.* 26, 6695–  
5 6703. <https://doi.org/10.1523/JNEUROSCI.0836-06.2006>
- 6 Driver, J., Noesselt, T., 2008. Multisensory Interplay Reveals Crossmodal Influences on  
7 “Sensory-Specific” Brain Regions, Neural Responses, and Judgments. *Neuron* 57, 11–23.  
8 <https://doi.org/10.1016/j.neuron.2007.12.013>
- 9 Failing, M., Theeuwes, J., 2018. Selection history: How reward modulates selectivity of visual  
10 attention. *Psychon. Bull. Rev.* 25, 514–538. <https://doi.org/10.3758/s13423-017-1380-y>
- 11 Failing, M.F., Theeuwes, J., 2014. Exogenous visual orienting by reward. *J. Vis.* 14, 6.  
12 <https://doi.org/10.1167/14.5.6>
- 13 Hickey, C., Chelazzi, L., Theeuwes, J., 2014. Reward-priming of location in visual search. *PLoS*  
14 *One* 9, e103372. <https://doi.org/10.1371/journal.pone.0103372>
- 15 Hickey, C., Chelazzi, L., Theeuwes, J., 2010. Reward changes salience in human vision via the  
16 anterior cingulate. *J. Neurosci.* 30, 11096–11103.  
17 <https://doi.org/10.1523/JNEUROSCI.1026-10.2010>
- 18 Hickey, C., Peelen, M. V., 2015. Neural mechanisms of incentive salience in naturalistic  
19 human vision. *Neuron* 85, 512–518. <https://doi.org/10.1016/j.neuron.2014.12.049>
- 20 Hickey, C., van Zoest, W., 2013. Reward-associated stimuli capture the eyes in spite of  
21 strategic attentional set. *Vision Res.* 92, 67–74.  
22 <https://doi.org/10.1016/j.visres.2013.09.008>
- 23 Hikosaka, O., Kim, H.F., Yasuda, M., Yamamoto, S., 2014. Basal Ganglia Circuits for Reward  
24 Value–Guided Behavior. *Annu. Rev. Neurosci.* 37, 289–306.  
25 <https://doi.org/10.1146/annurev-neuro-071013-013924>
- 26 Hillyard, S.A., Störmer, V.S., Feng, W., Martinez, A., McDonald, J.J., 2016. Cross-modal  
27 orienting of visual attention. *Neuropsychologia* 83, 170–8.  
28 <https://doi.org/10.1016/j.neuropsychologia.2015.06.003>
- 29 Ikeda, T., Hikosaka, O., 2007. Takuro Ikeda and Okihide Hikosaka Positive and Negative  
30 Modulation of Motor Response in Primate Superior Colliculus by Reward Expectation. *J*  
31 *Neurophysiol* 98, 3163–3170. <https://doi.org/10.1152/jn.00975.2007>
- 32 Johnson, J.A., Zatorre, R.J., 2005. Attention to simultaneous unrelated auditory and visual  
33 events: behavioral and neural correlates. *Cereb. Cortex* 15, 1609–1620.  
34 <https://doi.org/10.1093/cercor/bhi039>
- 35 Jolicoeur, P., 1999. Restricted attentional capacity between sensory modalities. *Psychon.*  
36 *Bull. Rev.* 6, 87–92.
- 37 Lucas, N., Schwartz, S., Leroy, R., Pavin, S., Diserens, K., Vuilleumier, P., 2013. Gambling  
38 against neglect: Unconscious spatial biases induced by reward reinforcement in healthy

- 1 people and brain-damaged patients. *Cortex* 49, 2616–2627.  
2 <https://doi.org/10.1016/j.cortex.2013.06.004>
- 3 Macaluso, E., 2010. Orienting of spatial attention and the interplay between the senses.  
4 *Cortex* 46, 282–297. <https://doi.org/10.1016/j.cortex.2009.05.010>
- 5 Macaluso, E., 2000. Modulation of Human Visual Cortex by Crossmodal Spatial Attention.  
6 *Science* (80-. ). 289, 1206–1208. <https://doi.org/10.1126/science.289.5482.1206>
- 7 Maclean, M.H., Giesbrecht, B., 2015. Neural evidence reveals the rapid effects of reward  
8 history on selective attention. *Brain Res.* 1606, 86–94.  
9 <https://doi.org/10.1016/j.brainres.2015.02.016>
- 10 McDonald, J.J., Green, J.J., Stormer, V.S., Hillyard, S.A., 2012. Cross-Modal Spatial Cueing of  
11 Attention Influences Visual Perception, in: Murray, M.M., Wallace, M.T. (Eds.), *The*  
12 *Neural Bases of Multisensory Processes*. Boca Raton (FL).
- 13 Peck, C.J., Jangraw, D.C., Suzuki, M., Efem, R., Gottlieb, J., 2009. Reward Modulates Attention  
14 Independently of Action Value in Posterior Parietal Cortex. *J. Neurosci.* 29, 11182–  
15 11191. <https://doi.org/10.1523/JNEUROSCI.1929-09.2009>
- 16 Pierrot-Deseilligny, C., Muri, R.M., Ploner, C.J., Gaymard, B., Rivaud-Pechoux, S., 2003.  
17 Cortical control of ocular saccades in humans: a model for motricity. *Prog. Brain Res.*  
18 142, 3–17. [https://doi.org/10.1016/S0079-6123\(03\)42003-7](https://doi.org/10.1016/S0079-6123(03)42003-7)
- 19 Pooremaeili, A., FitzGerald, T.H.B., Bach, D.R., Toelch, U., Ostendorf, F., Dolan, R.J., 2014.  
20 Cross-modal effects of value on perceptual acuity and stimulus encoding. *Proc. Natl.*  
21 *Acad. Sci. U. S. A.* 111, 15244–9. <https://doi.org/10.1073/pnas.1408873111>
- 22 Posner, M.I., 1980. Orienting of attention. *Q. J. Exp. Psychol.* 32, 3–25.
- 23 Qi, S., Zeng, Q., Ding, C., Li, H., 2013. Neural correlates of reward-driven attentional capture  
24 in visual search. *Brain Res.* 1532, 32–43. <https://doi.org/10.1016/j.brainres.2013.07.044>
- 25 Quentin, R., Chanes, L., Vernet, M., Valero-Cabré, A., 2015. Fronto-Parietal Anatomical  
26 Connections Influence the Modulation of Conscious Visual Perception by High-Beta  
27 Frontal Oscillatory Activity. *Cereb. Cortex* 25, 2095–2101.  
28 <https://doi.org/10.1093/cercor/bhu014>
- 29 Schneider, W., Eschman, A., Zuccolotto, A., Burgess, S., Cernicky, B., Gilkey, D., Gliptis, J.,  
30 Maciejczyk, V., Macwhinney, B., Rodgers, K., James, J.S., 2002. E-prime user's guide.  
31 *Pittsbg. Psychol. Softw. Tools Inc.*
- 32 Serences, J.T., 2008. Value-based modulations in human visual cortex. *Neuron* 60, 1169–  
33 1181. <https://doi.org/10.1016/j.neuron.2008.10.051>
- 34 Shomstein, S., Yantis, S., 2004. Control of attention shifts between vision and audition in  
35 human cortex. *J. Neurosci.* 24, 10702–10706. [https://doi.org/10.1523/JNEUROSCI.2939-](https://doi.org/10.1523/JNEUROSCI.2939-04.2004)  
36 04.2004
- 37 Spence, C., 2010. Crossmodal spatial attention. *Ann. N. Y. Acad. Sci.*  
38 <https://doi.org/10.1111/j.1749-6632.2010.05440.x>

- 1 Talsma, D., Senkowski, D., Soto-Faraco, S., Woldorff, M.G., 2010. The multifaceted interplay  
2 between attention and multisensory integration. *Trends Cogn. Sci.* 14, 400–410.  
3 <https://doi.org/10.1016/j.tics.2010.06.008>
- 4 Theeuwes, J., Belopolsky, A. V, 2012. Reward grabs the eye: Oculomotor capture by  
5 rewarding stimuli. *Vision Res.* 74, 80–85. <https://doi.org/10.1016/j.visres.2012.07.024>
- 6 van Koningsbruggen, M.G., Ficarella, S.C., Battelli, L., Hickey, C., 2016. Transcranial random-  
7 noise stimulation of visual cortex potentiates value-driven attentional capture. *Soc.*  
8 *Cogn. Affect. Neurosci.* 11, 1481–1488. <https://doi.org/10.1093/scan/nsw056>
- 9 Vroomen, J., De Gelder, B., 2004. Perceptual effects of cross-modal stimulation:  
10 Ventrioloquism and the freezing phenomenon. *Handb. multisensory Process.* 3, 1–23.
- 11 Vuilleumier, P., 2015. Affective and motivational control of vision. *Curr. Opin. Neurol.* 28,  
12 29–35. <https://doi.org/10.1097/WCO.0000000000000159>
- 13 Vuilleumier, P., Armony, J.L., Driver, J., Dolan, R.J., 2001. Effects of attention and emotion on  
14 face processing in the human brain: an event-related fMRI study. *Neuron* 30, 829–841.  
15 [https://doi.org/10.1016/S0896-6273\(01\)00328-2](https://doi.org/10.1016/S0896-6273(01)00328-2)
- 16 Vuilleumier, P., Brosch, T., 2009. Interactions of emotion and attention, in: Gazzaniga, M.S.  
17 (Ed.), *The Cognitive Neurosciences IV*. Cambridge: MIT Press, pp. 925–934.
- 18 Yasuda, M., Hikosaka, O., 2015. Functional territories in primate substantia nigra pars  
19 reticulata separately signaling stable and flexible values. *J. Neurophysiol.* 113, 1681–  
20 1696. <https://doi.org/10.1152/jn.00674.2014>

1 **Supplementary material**



2 **Figure S1.** Reward and trial blocks during the association phase. Mean saccadic RTs (ms)  
3 obtained during the association phase as a function of trial numbers and whether auditory  
4 stimuli were associated with a low (REW-) or high (REW+) reward.  
5

		Manual response		Total
		incorrect	correct	
Saccadic response	incorrect	14 (0.3)	363 (7.6)	377 (7.9)
	correct	54 (1.1)	4369 (91.0)	4423 (92.1)
Total		68 (1.4)	4732 (98.6)	4800 (100)

6  
7 **Table S1.** Accuracy of manual and saccadic responses. Number of trials and percentage of  
8 total trials (between brackets) during the testing phase with an incorrect or correct response  
9 using the manual response and the saccadic response modality.