Reviewed Preprint
v1 • April 15, 2026
Not revised

For correspondence:
perrine.porte@univ-grenoble-alpes.fr

* Shared contribution

Funding: See page 34

Reviewing editor: Jean-Paul Noel, University of Minnesota, United States

Audiovisual congruency drives confidence in presence and absence

Perrine Porte[1], Matan Mazor[2], Childéric Dezier[3], Nathan Faivre[1,*], Louise Goupil[1,*]

[1]Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, LPNC, Grenoble, France • [2]All Souls College and Department of Experimental Psychology, University of Oxford, Oxford, United Kingdom • [3]Univ. Grenoble Alpes, Inserm, U1216, Grenoble Institut Neurosciences, Grenoble, France

eLife Assessment

This **valuable** study investigates how multisensory signals influence detection decisions and confidence judgments in presence and absence tasks using pre-registered psychophysical experiments and computational modeling. Across two online samples, the authors argue that audiovisual stimuli improve detection performance but do not enhance metacognitive efficiency, and that confidence is higher for absence than presence judgments. The evidence is broadly **solid**, although aspects of the computational interpretation and model comparisons would benefit from additional clarification and testing against simpler alternatives.

https://doi.org/10.7554/eLife.110765.1.sa4

Abstract

The formation of a subjective sense of confidence often requires the integration of signals from multiple sources of evidence. This is of particular relevance when one needs to determine with a high degree of certainty whether a multisensory stimulus is present or absent. Understanding the mechanisms underlying this ability to map evidence strength from multiple modalities into a single confidence value is therefore central to the study of metacognition. To this end, we asked healthy adults to detect the presence or absence of near-threshold stimuli that could be visual, auditory, or both, and then rate their confidence. In two pre-registered experiments (N = 48 and N = 54), audiovisual stimuli were better detected than unimodal ones, but were not associated with better metacognitive performance. Surprisingly, participants were more confident in their absence than presence judgments. To explain these results, we fitted a Bayesian evidence accumulation model in which sensory evidence is available for presence only, rendering decisions about absence dependent on counterfactual inference. The model reproduced decision patterns by assuming that a stimulus was perceived if sensory evidence from either modality exceeded a threshold (a *disjunctive* integration rule). In contrast, it reproduced confidence judgments by assuming that high confidence requires that the two modalities align (*conjunctive* for presence, *disjunctive* for absence). Together, these findings reveal that distinct computational mechanisms drive perception and confidence when detecting near-threshold multisensory signals.

Introduction

Imagine being outside at night with soft music playing and dim light around you. If you faintly hear the buzz of a mosquito, are you more confident that it is present if you also catch a glimpse of it, and vice versa? Or is it sufficient for you to just see it – or to just hear it – to be confident that it is around? And if you neither see nor hear a mosquito, how can you be sure that no mosquito is around?

Using audiovisual detection tasks in the laboratory, it is now clearly established that human adults typically show higher detection rates and faster reaction times for bimodal stimuli than for unimodal stimuli (Rach et al., 2011 ; Plass & Brang, 2021). This behavior is well explained by evidence accumulation models, which assume that sensory evidence from each modality is accumulated and processed together until a decision threshold is reached, with decision times reflecting the time to reach the threshold (Gondan et al., 2004; Gondan et al., 2005; Plass & Brang, 2021; Egan et al., 2025). These models account for faster decision times in bimodal than in unimodal conditions, but provide limited insight into decision accuracy. To address this, Blurton et al. (2014) introduced a drift-diffusion model with two thresholds, enabling predictions of both decision times and accuracy in a go/no-go task in which participants had to detect a target stimulus and ignore distractors. Since a stimulus was always present in this task, no comparisons could be made between decisions about the presence versus absence of sensory evidence. In addition, these models did not explain people's ability to monitor the modality through which the stimulus was perceived.

If sensory evidence is only available for presence, a possibility is that absence is inferred from counterfactual perceptibility (i.e., *"I would have perceived a stimulus if it was present")* – a form of metacognitive belief (Mazor, 2025), with metacognition referring to the ability to monitor and control our own cognitive processes (Flavell, 1979). A recent modeling study showed that optimal visual detection relies on participants integrating factual sensory evidence for presence, and inferring absence from the absence of evidence (Mazor et al., 2025). This approach differs from previous evidence accumulation models, which typically assume that the same accumulation process similarly drives decisions about presence and absence. This Bayesian model was not only capable of explaining the features of detection responses (accuracy and response times) but also the confidence associated with these decisions, showing that confidence in absence relies on the use of "counterfactual evidence": beliefs about the evidence that would have been observed had a stimulus been present (Schipper & Mazor, 2025). This account helps clarify previously observed differences in metacognitive monitoring (i.e., the ability to make accurate confidence judgments about the results of our cognitive processes; Flavell, 1979; Fleming et al., 2012) between detection and discrimination tasks. Metacognitive judgments have been found to be more precise after discrimination than detection tasks, and this seems to be explained by the specific nature of absence judgments, which may rely on higher-order inferential processes (Meuwese et al., 2014; Kellij et al., 2020; Mazor et al., 2020; Mazor & Fleming, 2020; Mazor et al., 2023). However, in everyday life, our percepts are essentially multisensory, so a realistic model of detection – and confidence in these decisions – should also be able to explain how people judge the presence and absence of multisensory objects. This is a non-trivial problem, because in a multisensory context, observers have to combine sensory channels in order to decide if something is present or not, and such a combination can be achieved in several ways. Furthermore, it is unclear whether the same integration rules apply to detection and confidence.

We still have limited insight into these questions because how multisensory information impacts confidence judgments is only a recent domain of interest (Deroy et al., 2016). Recent studies comparing bimodal and unimodal conditions showed that, despite improvements in task performance, metacognitive efficiency was not improved in multisensory compared to unisensory conditions (Charles et al., 2020; Arbuzova et al., 2021; Faivre et al., 2018). Crucially, however, all these studies have used discrimination tasks, leaving the impact of a multisensory context on confidence in absence unknown.

Here, our goal was to understand how confidence judgments about the presence or absence of a stimulus are formed when evidence is available across multiple sensory channels. To this end, in two preregistered experiments, participants judged whether a stimulus was present or absent, irrespective of the modality, and rated their confidence in this amodal decision before providing modality-specific judgments and confidence ratings. Based on the literature relying on discrimination tasks, we hypothesized that multisensory stimuli would be associated with improved detection performance, but not metacognitive efficiency (Charles et al., 2020; Arbuzova et al., 2021; Faivre et al., 2018). We also expected participants to be more confident and have a higher metacognitive efficiency for presence than for absence judgments (Mazor et al., 2020; Mazor & Fleming, 2020; Meuwese et al., 2014; Kellij et al., 2020). We assessed whether factual and counterfactual reasoning applied to audiovisual detection by extending the model of Mazor and collaborators (2025) to the audiovisual domain. Finally, we examined the audiovisual integration rules governing confidence in presence and absence by deriving confidence from the probability of being correct at the time of the decision. Together, our behavioral and modeling results support the view that inferring audiovisual presence relies on a combination of factual and counterfactual reasoning, with distinct integration rules across sensory modalities for detection and confidence judgments.

Results

In two pre-registered online experiments (**Exp.1**: *https://osf.io/3nvyx*, **Exp.2**: *https://osf.io/ehndv*), participants performed an audiovisual detection task in which the stimulus could be either only visual, only auditory, or audiovisual. At the beginning of the experiment, stimulus intensity was calibrated for each participant to reach a 50% detection rate in visual and auditory conditions. Participants indicated whether a stimulus was present or absent irrespective of its sensory modality, before reporting their amodal confidence in their detection choice on a scale from 0 (“sure incorrect”) to 100 (“sure correct”). Finally, they reported their modality-specific detection and confidence judgments on a bi-dimensional (audio/visual) report scale, with each axis ranging from 100% sure not perceived to 100% sure perceived, and corresponding to one modality (see Fig.1). In Experiment 1, each experimental condition was equiprobable; in Experiment 2, we increased the proportion of trials without any signal to obtain an equal proportion of target absent and target present trials. Data from 48 participants were included in the main analysis for Experiment 1, and 54 for Experiment 2.

Response bias towards absence

In both experiments, participants demonstrated good task sensitivity (**Exp.1**: *d’* = 1.59 (*SD* = 0.64); **Exp.2:** *d’* = 1.77 (*SD* = 0.61)). They also had a significant bias toward responding “absent”, with a Bayesian one sample t-test showing a significant positive response criterion (**Exp.1**: *c* = 0.58, *SD* = 0.43, *BF1* > 1000; **Exp.2**: *c* = 0.43, *SD* = 0.48, *BF1* > 1000) (see Fig.2A).

Improved detection for multisensory stimuli

In both experiments, Bayesian logistic regression analyses on amodal accuracy revealed that participants were better at detecting bimodal than unimodal stimuli (**Exp.1**: $\hat{\beta}$ = 0.38, 95% CI [0.33,0.42], *BF1* > 1000; **Exp.2**: $\hat{\beta}$ = 0.41 [0.36,0.46], *BF1* > 1000). Despite our attempt to equalize performance between unimodal conditions, stimuli were better detected in the visual than in the auditory modality (**Exp.1**: $\hat{\beta}$ = 0.38 [0.17,0.59], BF1 = 38; **Exp.2**: $\hat{\beta}$ = 0. 56 [0. 38,0.74], BF1 > 1000) (see Fig.2B).

Ability to monitor the source of the percept

Our bi-dimensional report scale allowed us to assess how accurately participants identified the source of their percepts. Focusing on correct presence judgments, we performed exploratory analyses to investigate the ability of participants to accurately categorize the modality in which the stimulus was presented — a form of source monitoring (“*did I see or hear this stimulus?*”) Participants demonstrated reliable source monitoring: auditory stimuli were correctly categorized as auditory on 85% of trials in Experiment 1 (binomial test against 50%, *p* < .001) and 79% in Experiment 2 (*p* < .001), while visual stimuli were correctly categorized as visual on 75% (*p* < .001) and 82% of trials (*p* < .001), respectively.

For audiovisual stimuli, participants categorized them as auditory in 27% (**Exp.2:** 17%), visual in 39% (**Exp.2**: 45%), and audiovisual in 33% (**Exp.2**: 37%) of trials. These proportions are consistent with the idea that detecting stimuli as audiovisual is based on the joint probability of detecting them visually and auditorily. A Wilcoxon signed-rank test comparing observed audiovisual reports to the predicted joint probability revealed no significant difference (**Exp.1:** *V* = 764, *p* = .07, **Exp.2:** *V* =811, *p* = .56). This indicates that audiovisual categorizations do not violate the predictions from

Figure 1. **Trial structure.**

1) Amodal detection: Participants were asked to indicate if they perceived or not a stimulus, irrespective of its modality. In frame: different possible stimulus types: the auditory stimulus was a sinusoidal tone of 1kHz presented in pink noise, the visual stimulus a light gray circle presented in dynamic Gaussian noise. 2) Amodal confidence: Participants indicated their confidence in their detection answer. 3) Modality-specific judgments and confidence: By moving a cursor on a bi-dimensional scale, participants indicated simultaneously whether they perceived a stimulus in each modality, and with which level of confidence.

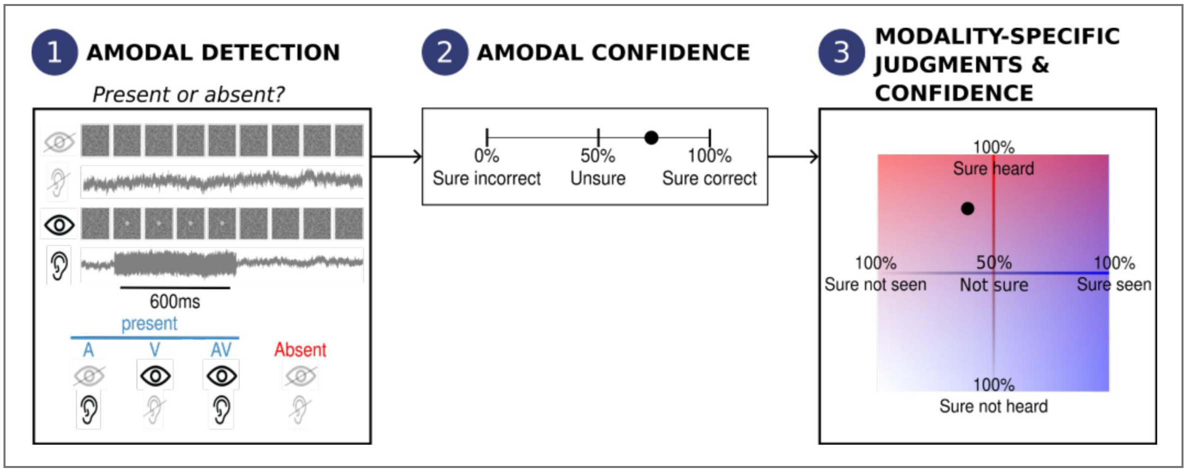
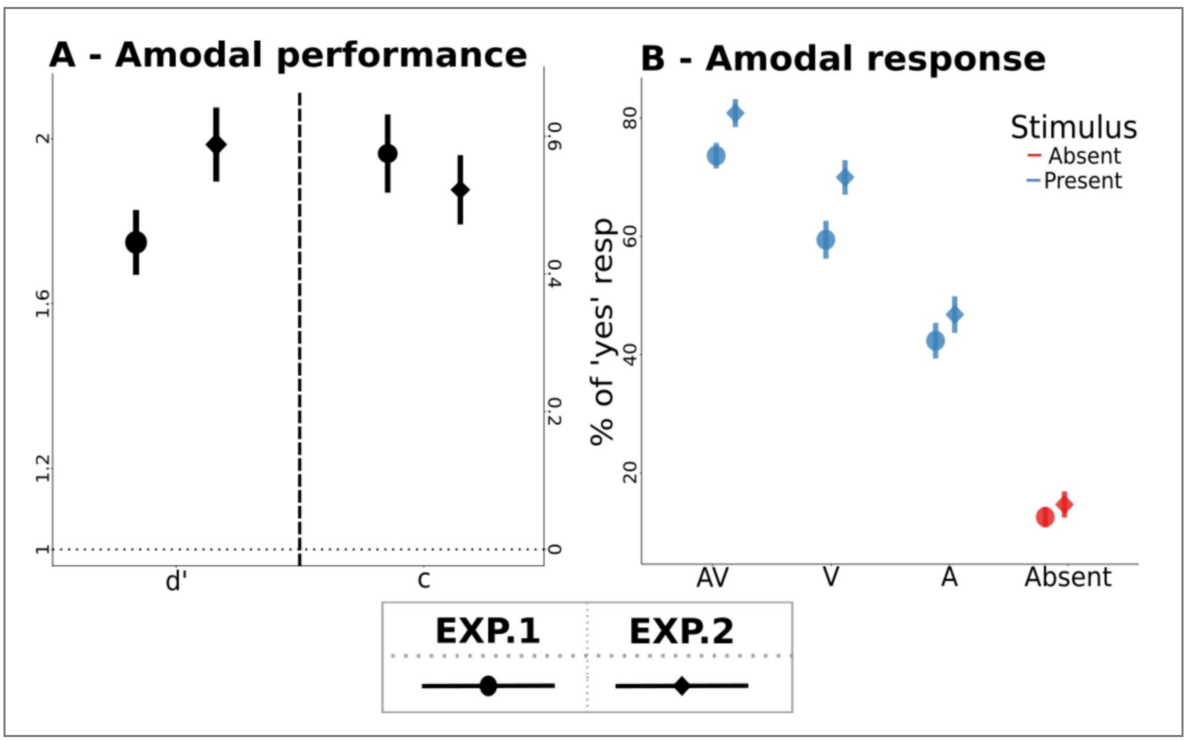


Figure 2. **Amodal detection.**

Experiment 1 is represented by circles, while Experiment 2 is represented by diamonds; error bars represent the standard error. A) Amodal performance: amodal d' and criterion. B) Amodal response: Percentage of stimuli judged to be present as a function of the experimental condition.



independent detection in the visual and auditory modalities. Finally, participants did not change their minds and attributed perceived presence to at least one modality after indicating that the stimulus was present, with only 1% of detected stimuli judged as being absent in both modalities.

Perceptual multisensory interference

Our bimodal confidence scale also enabled us to investigate multisensory interference (i.e., how does one modality influence the other). We found that the presence of one modality biased participants toward reporting that a stimulus was also present in the other, especially when it was, in fact, not the case. Indeed, the presence of a visual stimulus biased participants to respond that an auditory stimulus was present only when the auditory stimulus was actually absent. Likewise, the presence of auditory stimulus also biased participants to respond that a visual stimulus was present only when it was actually absent (see Table 1 and SI).

Table 1. Effect of the presence of a stimulus in the other modality

		Detection of absence	Detection of presence	Metasensitivity for absence	Metasensitivity for presence	Metaefficiency of absence judgments	Metaefficiency of presence judgments
Influence of visual evidence on auditory responses	**EXP.1**	*Reduced*	*No effect*	***Reduced***	*No effect*	*Reduced*	*No effect*
	EXP.2	*No effect*	*No effect*	***Reduced***	*Improved*	*No effect*	*No effect*
Influence of auditory evidence on visual responses	**EXP.1**	*No effect*	*No effect*	*No effect*	***Improved***	*No effect*	*No effect*
	EXP.2	*Reduced*	*No effect*	*Reduced*	***Improved***	*No effect*	*No effect*

Note: Bold indicates replicated effect across experiments. See SI for detailed results.

Confidence in presence vs absence

Based on previous findings (Mazor et al., 2020; Mazor et al., 2025), we hypothesized that participants would be more confident in their presence than in their absence judgments. Contrary to our hypothesis, in both experiments, participants reported higher amodal confidence following absence than presence judgments (**Exp.1**: $\hat{\beta} = -4.2$ [−5.6,−2.78], BF1 > 1000; **Exp.2**: $\hat{\beta} = -3.78$ [−4.96,−2.64], BF1 > 1000) (see Fig.3A).

To evaluate metacognitive performance irrespective of task performance (Fleming & Lau, 2014, Maniscalco & Lau, 2012), we analyzed metacognitive efficiency (Mratio, estimated via response-conditional HMeta-d; Fleming, 2017) for presence versus absence judgments, both at the amodal and at the modality-specific level. In both experiments, and consistent with previous findings (Mazor et al., 2020; Mazor & Fleming, 2020), amodal metacognitive efficiency was higher for presence judgments than absence judgments (**Exp.1**: ΔM = 0.53 [0.27, 0.79]; **Exp.2**: ΔM = 0.62 [0.43, 0.81]) (see Fig.3B). Using the bidimensional scale, we found that this effect was also true at the modality-specific level: participants showed higher auditory metacognitive efficiency in auditory judged present compared to auditory judged absent trials (**Exp.1**: ΔM = 0.78 [0.53, 1.03]; **Exp.2**: ΔM = 0.67 [0.41, 0.93]), and higher visual metacognitive efficiency in visually judged present compared to visually judged absent trials (**Exp.1**: ΔM = 0.60 [0.40, 0.81]; **Exp.2**: ΔM = 0.35 [0.14, 0.56]).

Multisensory effects on metacognitive performance

To evaluate how accurately participants' confidence tracked their accuracy, we measured metacognitive sensitivity as the slope of the Bayesian logistic regression predicting accuracy from confidence. Amodal metacognitive sensitivity was higher in bimodal than in unimodal trials (**Exp.1:** $\hat{\beta}$=0.11 [0.07,0.15], BF1 > 1000; **Exp.2:** $\hat{\beta}$= 0.11 [0.07,0.15], BF1 > 1000). Experiment 1 showed moderate evidence for an absence of a difference in amodal metacognitive sensitivity between auditory and visual conditions ($\hat{\beta}$ = 0.04 [−0.06,0.14], BF0 = 5.38), but the evidence was inconclusive in Experiment 2 (BF0 = 0.62) (see Fig.4A).

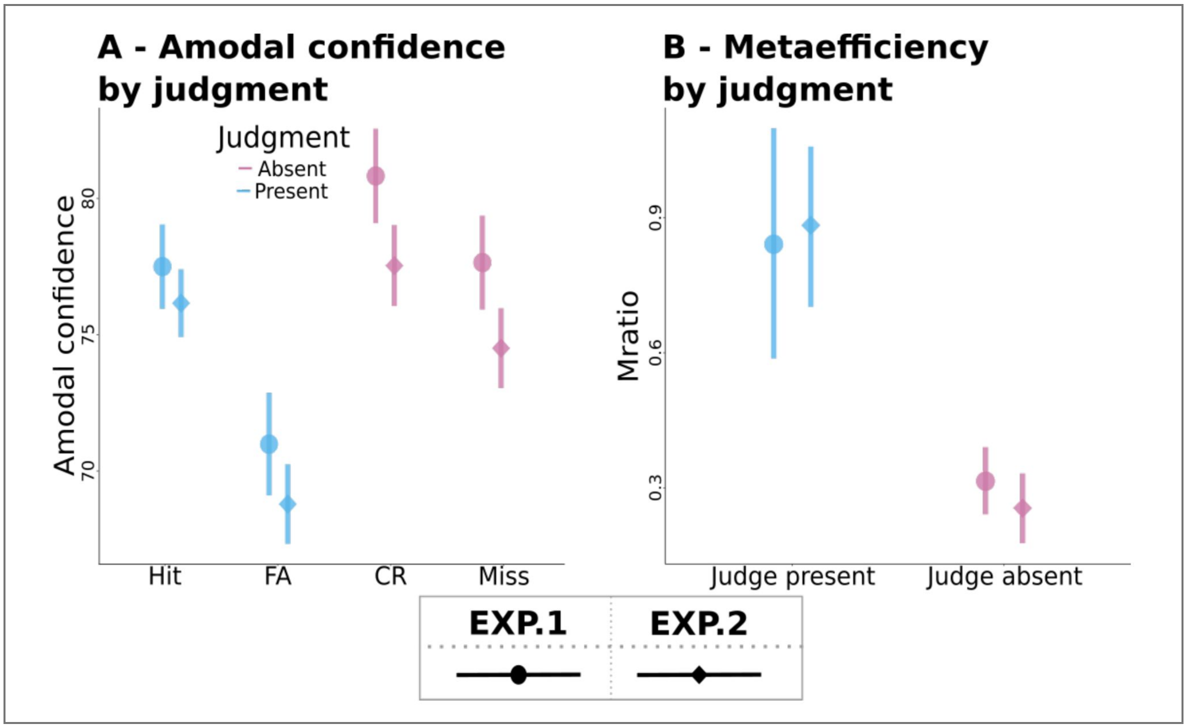


Figure 3. **Confidence by judgment.**

Experiment 1 is represented by circles, while Experiment 2 is represented by diamonds. A) Amodal confidence as a function of hits, false alarms (FA), correct rejections (CR), and misses; error bars represent the standard error. B) Amodal metacognitive efficiency (response-conditional Mratio) as a function of the type of judgments; error bars represent the highest density interval. In both panels, presence judgements are represented in blue, and absence judgements in pink.

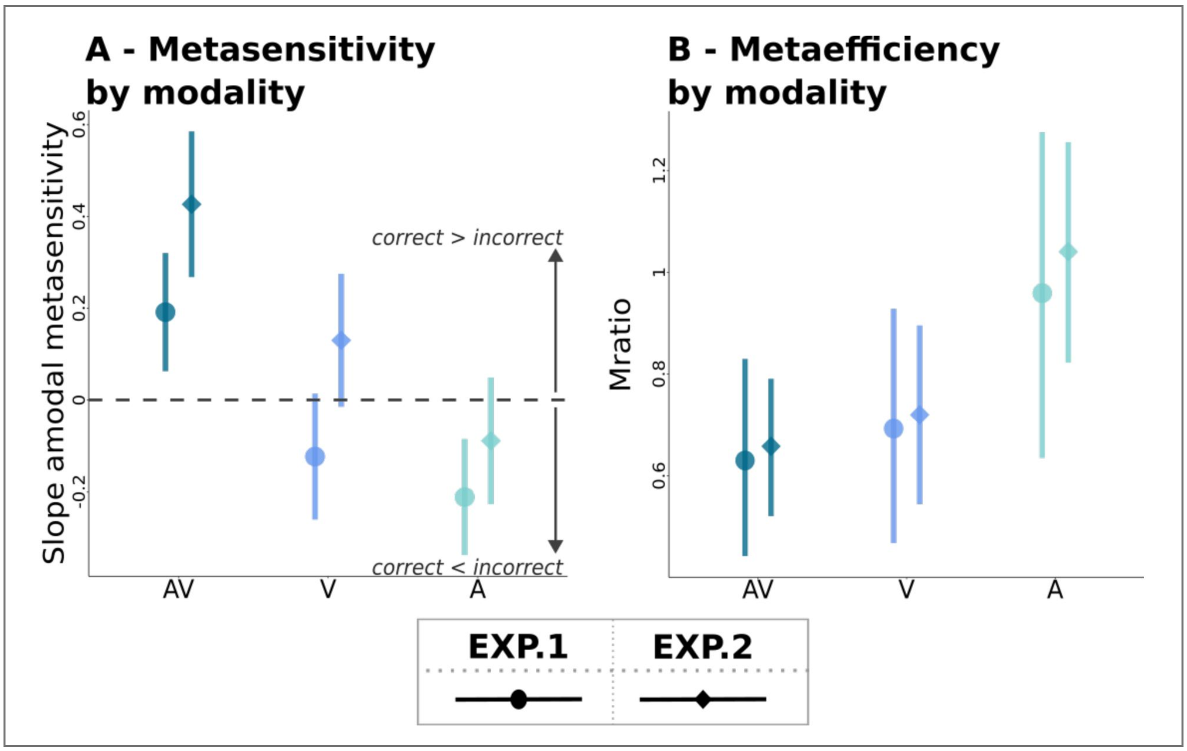


Figure 4. Multisensory effects.

Experiment 1 is represented by circles, while Experiment 2 is represented by diamonds. A) Amodal metacognitive sensitivity by modality as a function of the experimental condition, a slope superior to 0 indicates higher confidence in correct than incorrect responses and below 0 higher confidence in incorrect than correct responses; error bars represent the standard error. B) Amodal metacognitive efficiency (response-conditional Mratio) for presence judgments as a function of the modality of presentation; error bars represent the highest density interval.

Additionally, we analyzed metacognitive efficiency across stimulus modalities (see Fig.4B). Focusing on presence judgments, in Experiment 1, no credible difference was found between unimodal and bimodal trials (ΔM = -0.25 [-0.59, 0.10]). There was also no clear difference in metacognitive efficiency between auditory and visual trials (ΔM = 0.27 [-0.15, 0.69]), between auditory and audiovisual trials (ΔM = -0.30 [-0.67, 0.07]), or between visual and audiovisual trials (ΔM = -0.03 [-0.29, 0.26]). In contrast, in Experiment 2, we observed higher metacognitive efficiency in unimodal compared to bimodal trials (ΔM = -0.28 [-0.54, -0.02]). This effect was driven by auditory trials with both a higher metacognitive efficiency for auditory compared to visual trials (ΔM = 0.32 [0.03, 0.59]) and for auditory compared to audiovisual trials (ΔM = -0.38 [-0.63, -0.13], while there was no evidence for a difference in metacognitive efficiency between visual and audiovisual trials (ΔM = -0.06 [-0.29, 0.16]). Finally, we found strong evidence for the existence of correlations between metacognitive efficiencies in the audiovisual and auditory domain (**Exp.1**: *rho* = 0.86 [0.61, 0.99]; **Exp.2:** *rho* = 0.78 [0.42, 0.99]), between the audiovisual and visual domain (**Exp.1:** *rho* = 0.89 [0.72, 0.99]; **Exp.2:** *rho* = 0.92 [0.51, 0.99]), and between the auditory and visual domain (**Exp.1**: *rho* = 0.85 [0.61, 0.99]; **Exp.2:** *rho* = 0.76 [0.25, 0.98]), consistent with other findings on the supramodality of metacognition (Rouault et al., 2018; Faivre et al., 2018; Ais et al., 2016).

Metacognitive multisensory interference

As preregistered, we also examined multisensory interference, this time at the metacognitive level. Our results suggest a cross-modal facilitation when both modalities provide consistent cues: in both experiments, participants showed higher auditory metacognitive sensitivity for absence when both the visual and the auditory stimuli were absent and higher visual metacognitive sensitivity for presence when both the visual and the auditory stimuli were present. We also observed additional but less robust influence, present in only one of the two experiments; given this lack of reproducibility, we do not interpret these effects further (see Table 1 and SI).

Model

We extended a recent ideal-observer model of visual detection to account for our multisensory detection task (Mazor et al., 2025). In the original model, observers determine the presence of a visual object depending on the activation of a single “presence sensor”. At each time point, this sensor samples either a 1 or a 0, and the probability of sampling a 1 is higher when the target is present. Sensor activation probabilities are captured by model parameters $\theta_{present}$ (i.e., probability of sampling 1 when a target is present) and θ_{absent} (i.e., probability of sampling 1 when a target is absent). Importantly, this model assumes that agents hold beliefs about these probabilities (i.e., beliefs about the probability of sampling a 1, if the target is present or if it is absent). These beliefs are captured by model parameters $\bar{\theta}_{present}$ and $\bar{\theta}_{absent}$, and reflect the degree to which agents believe that they would have perceived the target if it were present. After each sample, the agent decides whether to respond “present”, respond “absent”, or wait and accumulate more evidence. This decision is made based on the optimal policy (derived using backward induction; Callaway et al., 2023), assuming temporal discounting of the value of being correct, and based on the log-likelihood ratio between the two competing hypotheses (present or absent). A detailed technical description of the model is available in Mazor et al. (2025).

To adapt this model for audiovisual detection, we implemented two modality-specific sensors: a visual and an auditory sensor (see Fig.5). The model also allowed holding distinct prior beliefs regarding the probability of presence in each modality. At each time point, the posterior probability of a stimulus being present was updated independently in each sensory channel. Because participants were instructed to respond whenever they detected a stimulus, regardless of the sensory modality, we combined the two channels into an amodal estimate using a disjunctive integration rule, such that $p(present)=p(A_{present})+p(V_{present})-p(A_{present}\ and\ V_{present})$, reflecting that a stimulus can be (objectively) present in only one modality (A, V) or in both modalities (AV). We compared it to a conjunctive integration rule according to which a stimulus was judged present only if present in both modalities (see SI for comparison).

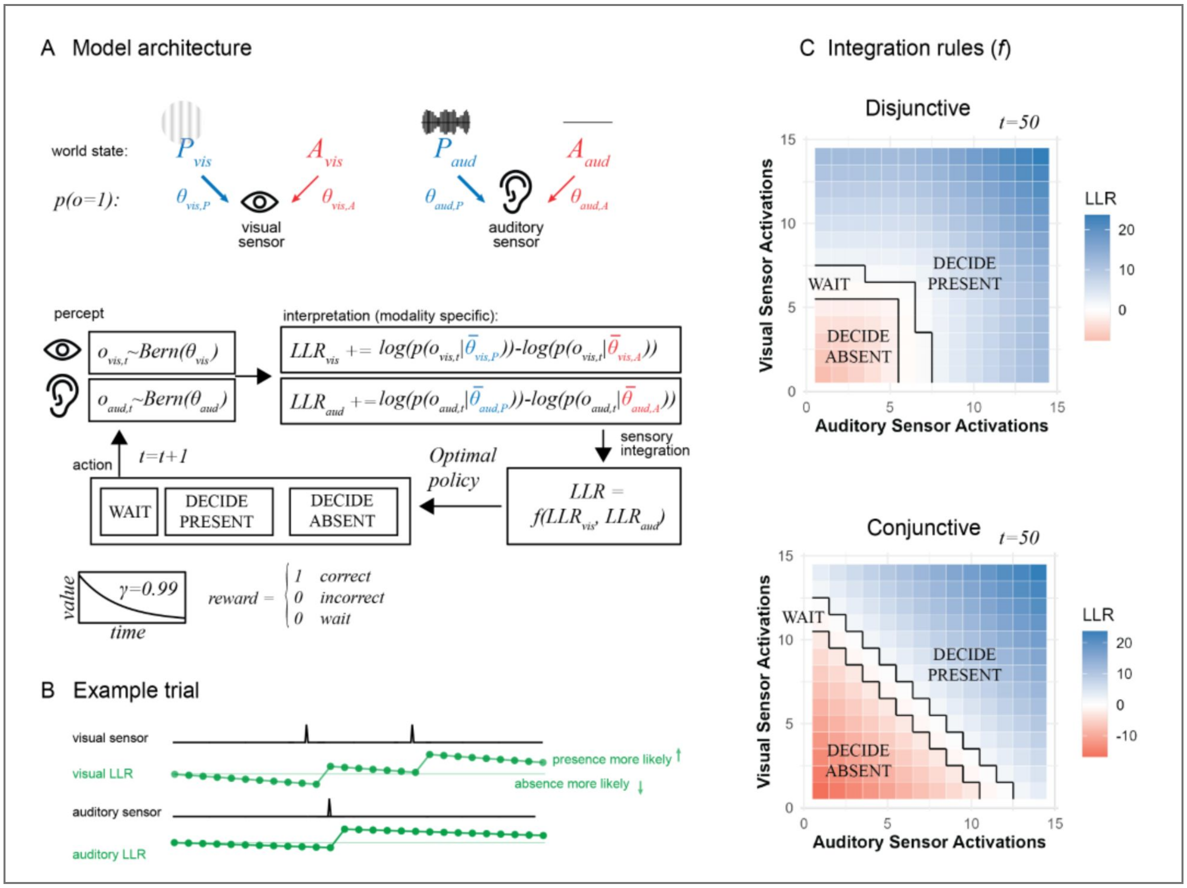


Figure 5. Computational model.

A) Model architecture. The observer is assumed to have access to a visual sensor and an auditory sensor, probabilistically tuned to the presence of visual and auditory evidence. The probability of activation is controlled by the parameter θ. The model agent observes the activations and updates their beliefs about the presence of a signal in each modality separately, using Bayes' rule. The agent then integrates the two beliefs into an amodal belief in the presence of a target. Based on this belief, they decide whether to commit to a decision or accumulate more evidence by following an optimal policy, derived using backward induction. B) Example trial: modality-specific log-likelihood ratios (LLR, in green) are updated following sensor inactivations and activations. C) Integration rules: The top plot represents the disjunctive rule and the bottom plot the conjunctive rule. Amodal LLR is plotted as a function of the number of sensor activations in each modality 50 time points (i.e., 2.5s) into the trial. Black contours indicate regions in which the best action is to decide present, wait, or decide absent.

Although the model was only fitted to amodal detection accuracy and response time data, it generated qualitative predictions about amodal confidence based on the probability of being correct at the time of the decision. Furthermore, it made predictions about modality-specific effects as it computed the probability of presence separately for each modality: a stimulus was judged present if the modality-specific probability of presence was greater than 0.5 at the time of the decision, and absent otherwise. Modality-specific confidence was also read as the probability of being correct at the time of the decision, for each modality separately.

To assess the model's ability to reproduce qualitative patterns observed in our behavioral data, we simulated datasets using a model in which belief parameters $\bar{\theta}$ are shared across modalities (hereafter, single-belief model; see Method for model comparisons). Specifically, we simulated data from 48 participants using parameters fitted to individual participants from Experiment 1, and from 54 participants using parameters fitted to individual participants from Experiment 2.

Reproduction of perceptual effects

The model successfully reproduced amodal detection behavior, with a mean d' of 1.67 (*SD* = 0.86) and a mean criterion of 0.59 (*SD* = 0.49) for Experiment 1, and a mean d' of 1.73 (*SD* = 0.86) and a mean criterion of 0.46 (*SD* = 0.47) for Experiment 2. It also reproduced higher accuracy for audiovisual trials than for unimodal trials, and for visual trials than for auditory trials in both experiments (see Fig.6A; Fig.6D; fits for the response times are shown in Fig.S6

Despite being trained only to amodal detection responses, the model also reproduced the observed modality-specific detection patterns in both experiments (see Fig.6B-C; Fig.6E-F). It also reproduced participants' capacity for accurate source monitoring (see Fig.6G-J). Notably, only 5% of stimuli judged to be present were modeled as absent in both modalities, suggesting that the model captured participants' response consistency by attributing perceived presence to at least one modality.

Reproduction of confidence effects

Having shown that optimal amodal detection involved integrating sensory evidence according to a disjunctive rule (i.e., a stimulus is detected based on auditory *or* visual evidence), we sought to test whether the same rule applied for amodal confidence. To do so, we tested whether, at the time of the decision, amodal confidence was based on the probability of being correct according to a disjunctive or conjunctive integration of information, separately for presence and absence judgments. If amodal confidence follows a disjunctive integration rule, confidence in presence should be high when only one modality indicates presence, whereas confidence in absence should be high only when both modalities indicate absence (i.e., negation of the disjunction). On the other hand, if amodal confidence follows a conjunctive integration rule, confidence in presence should be high only when both modalities indicate presence, whereas confidence in absence should be high when only one modality indicates absence (i.e., negation of the conjunction).

When confidence followed a disjunctive rule, the model failed to capture important aspects of the data, such as the higher confidence for presence than absence judgments (Fig.7, left panels). On the other hand, when confidence followed a conjunctive rule, it reproduced confidence in presence judgments but failed to capture variability in confidence ratings for absence judgments (Fig. 7, middle panels). Critically, a combination of disjunctive and conjunctive integration rules, for absence and presence judgments respectively, reproduced the confidence effects we observed in both experiments (while being generally over-confident relative to human participants across the board). Namely, it predicted higher confidence in absence than in presence judgments, higher metacognitive sensitivity for bimodal compared to unimodal trials, and higher confidence for correct than for incorrect absence judgments (see Fig.7, right panels). This combined rule can be alternatively described in the following way: high confidence, whether in presence or in absence, requires that the two sensors agree: both indicate the presence of a signal for high confidence "presence" judgments, or that both indicate signal absence for high confidence "absence" judgments.

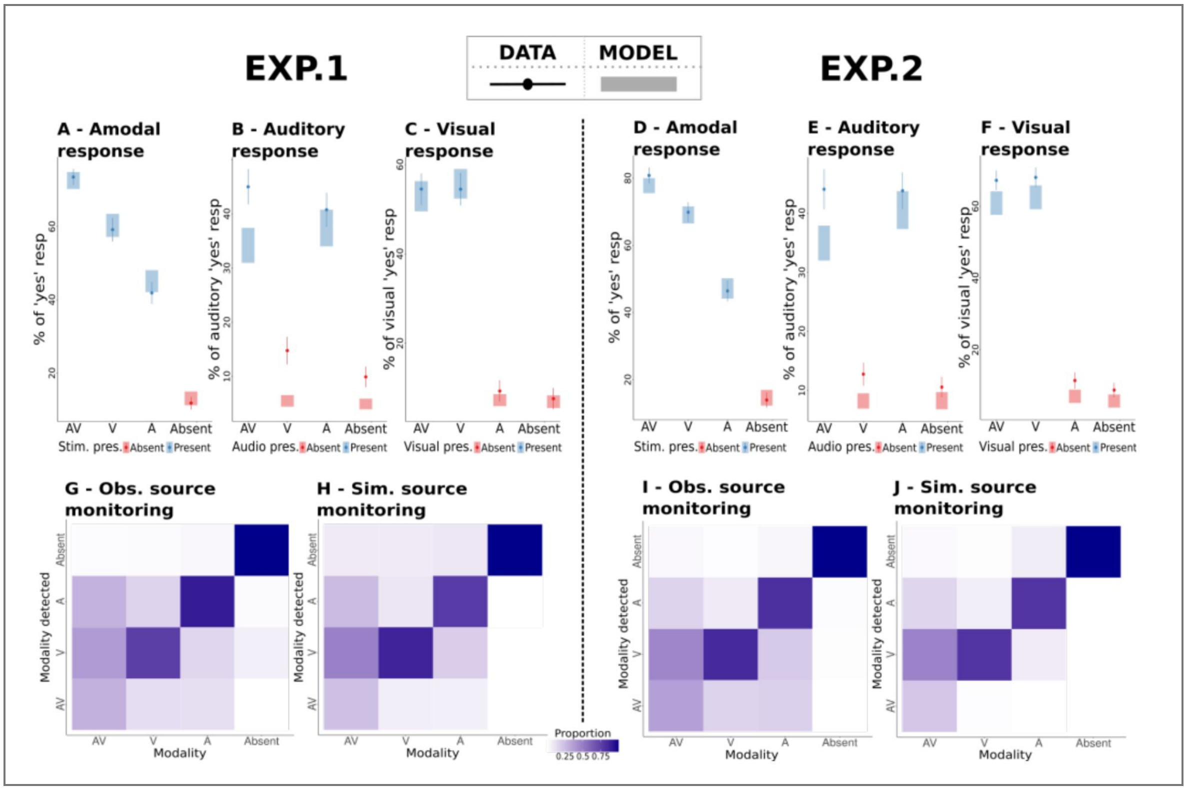


Figure 6. **Reproduction of perceptual effects.**

Error bars represent the standard error from the data. Rectangles represent data simulated from the model, centered on the mean value and with height equal to the standard error. Left panels show the fit for Experiment 1 and right panels for Experiment 2. A-F) Percentage of stimuli judged to be present as a function of the condition of presentation for Experiments 1 and 2, at the amodal, auditory, or visual level. G-J) Observed and simulated source monitoring: Modality detected as a function of the modality of presentation in Experiments 1 and 2.

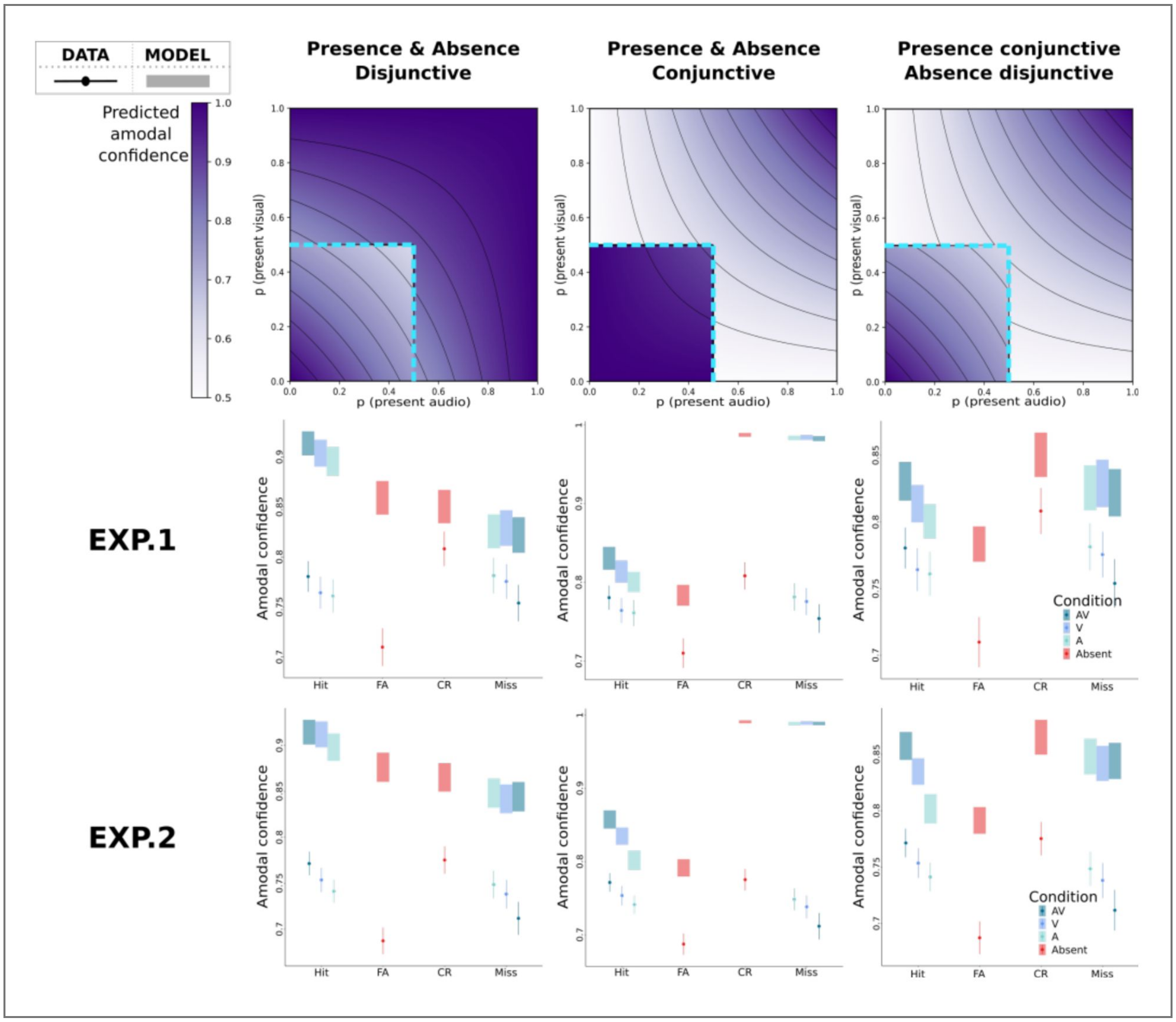


Figure 7. **Confidence fits according to the different integration rules.**

Error bars represent the standard error from the data. Rectangles represent data simulated from the model, centered on the mean value and with height equal to the standard error. Top plots represent the fit for Experiment 1 and bottom plots for Experiment 2. Each plot represents the amodal confidence as a function of condition of presentation and as a function of amodal hits, false alarms (FA), correct rejections (CR), and misses. The left panels represent the confidence based on the disjunctive rule. The middle panels represent the confidence based on the conjunctive rule. The right panels represent the confidence when absence is based on the disjunctive rule, while presence is based on the conjunctive rule.

Despite being fitted to amodal decisions only, the model captured the observed modality-specific confidence effects by reproducing confidence as a function of participants' response and accuracy both in the visual and in the auditory modality (see Fig.8). Finally, we tested if the model captured interindividual variability in confidence asymmetries between the auditory and visual modalities, reflecting a propension to give more weight to one sensor when estimating confidence following an audiovisual stimulus. To do so, we defined for each participant a "confidence asymmetry index" capturing the difference between auditory and visual confidence in audiovisual trials, normalised by absent trials:

$$CAI = ([auditory\, confidence]_{AV} - [auditory\, confidence]_{Absent}) \\ - ([visual\, confidence]_{AV} - [visual\, confidence]_{Absent})$$

A significant positive correlation was observed between the observed and simulated confidence asymmetry indices (**Exp.1:** *Pearson's rho* = .84, $p < .001$; **Exp.2**: *Pearson's rho* = .61, $p < .001$), indicating that the model successfully captured the interindividual variability of visual and auditory weights on confidence.

Discussion

Although everyday perception is inherently multisensory, we know surprisingly little about the way people judge whether something is present or absent across multiple sensory channels, and how confident they are in such judgments. To address this, in two preregistered experiments, participants performed an audiovisual task at unimodal near-threshold intensity. On each trial, they reported whether a stimulus was present irrespective of the modality of presentation before reporting their amodal confidence in their answer, and finally, their modality-specific judgments and confidence.

To investigate audiovisual integration rules, we adapted a recent Bayesian evidence accumulation framework assuming that absence is inferred from counterfactual detectability (Mazor et al., 2025). Using a disjunctive integration rule according to which a stimulus is detected when at least one modality provides sufficient evidence, we reproduced the higher detection performance for audiovisual stimuli compared to unimodal ones. Importantly, our model showed that audiovisual signals were processed differently at the perceptual and metacognitive levels. The model successfully reproduced the observed amodal confidence patterns using a disjunctive rule for absence judgments (i.e., high confidence in absence only when neither modality provided sufficient evidence), but using a conjunctive rule for presence judgments (i.e., high confidence in presence only when both modalities provided sufficient evidence). This shows that high confidence in absence or presence occurs when the auditory and the visual channels are aligned. This suggests that, while detection decisions relied on a disjunctive process, where evidence from a single modality is sufficient to judge that something is present, at the confidence level, intersensory congruency played a critical role. This interpretation is further supported by the modality-specific confidence effects we observed, which indicated cross-modal facilitation at the metacognitive level when both modalities provided consistent cues.

Looking more closely at multisensory effects, we replicated previous findings showing higher detection performance, and, for correct responses only, faster response times for audiovisual compared to unimodal stimuli (Rach et al., 2011; Plass & Brang, 2021). Using a bidimensional modality-specific scale, we also observed that participants accurately monitor the source of their percepts, by identifying the modality in which the stimulus was presented. At the metacognitive level, despite an improvement in metacognitive sensitivity, there was no boost in metacognitive efficiency of presence judgments for bimodal compared to unimodal stimuli. Experiment 2 even showed a higher metacognitive efficiency for auditory trials; given that this effect did not replicate across experiments, further work is needed to test its reliability. Prior work has similarly reported no differences in metacognitive efficiency between bimodal and unimodal stimuli during discrimination tasks (Charles et al., 2020; Arbuzova et al., 2021; Faivre et al., 2018), and our findings extend these results to detection tasks. Furthermore, we found strong correlations in

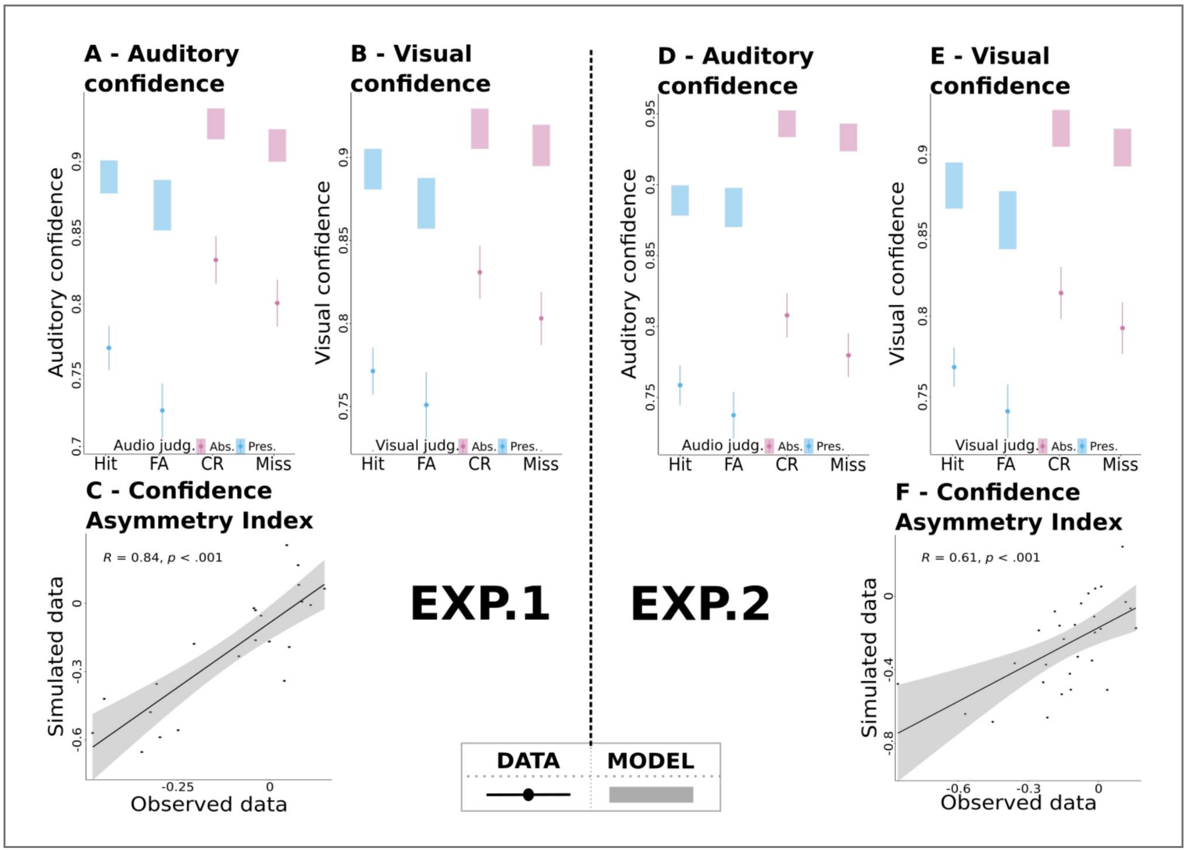


Figure 8. **Reproduction of modality-specific confidence effects.**

Error bars represent the standard error from the data. Rectangles represent data simulated from the model, centered on the mean value and with height equal to the standard error. Left panels represent the fit for Experiment 1 and right panels for Experiment 2. A) Auditory confidence as a function of auditory hits, false alarms (FA), correct rejections (CR), and misses for Experiment 1. B) Visual confidence as a function of visual hits, false alarms (FA), correct rejections (CR), and misses for Experiment 1. C) Correlation between observed and simulated data for the confidence asymmetry index for Experiment 1. D) Auditory confidence as a function of auditory hits, false alarms (FA), correct rejections (CR), and misses for Experiment 2. E) Visual confidence as a function of visual hits, false alarms (FA), correct rejections (CR), and misses for Experiment 2. F) Correlation between observed and simulated data for the confidence asymmetry index for Experiment 2. In all panels, presence judgements are represented in blue, and absence judgements in pink.

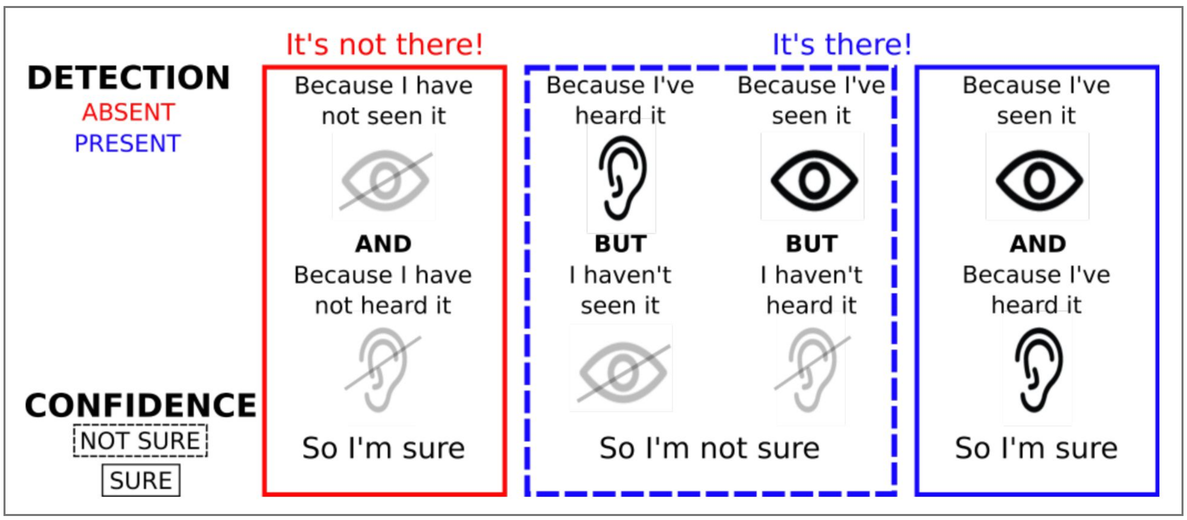


Figure 9. Illustration of the integration rules process.

Detection decisions (red for absence, blue for presence) are based on the disjunctive integration rule (disjunction and negation of disjunction). Confidence decisions (dashed line for not sure, full line for sure) are either based on a conjunctive rule (confidence in presence) or a negation of disjunction (confidence in absence).

metacognitive efficiency across modalities, consistent with findings suggesting the supramodality of metacognition (Rouault et al., 2018; Faivre et al., 2018; Ais et al., 2016). Finally, the modality-specific scale we developed provides a promising tool for investigating perceptual and metacognitive asymmetries in the presence or absence of multisensory stimuli, particularly in populations with sensory impairments, who may differ in both factual and counterfactual reasoning.

When investigating the metacognitive monitoring of absence, we found that participants were overall more confident in their absence than in their presence judgments, both at the amodal and modality-specific levels. This is in contradiction with some previous findings (Meuwese et al., 2014; Kellij et al., 2020; Mazor et al., 2025), but consistent with other studies (Pereira et al., 2021; Stockart et al., 2025; Dijkstra et al., 2024). As there are multiple differences between these experimental paradigms, it is difficult to pinpoint what could drive higher confidence in absence. To examine whether stimulus intensity contributes to this confidence pattern, we performed an additional experiment in which unimodal stimuli were presented at a suprathreshold intensity (see SI for detailed results). In this configuration, participants were still more confident in their absence than in their presence judgments, suggesting that increased sensory evidence is not sufficient to restore high confidence in presence. Future research will be needed to further investigate the role of the multisensory context on this effect. Indeed, if absence judgments rely on counterfactual reasoning, one can hypothesize that when multiple sensory sources are available but no stimulus is perceived in any of them, the belief that you should have perceived the stimulus if it was present could become even stronger, potentially increasing confidence in absence. Nevertheless, in all our experiments, despite higher confidence in absence, participants had a higher metacognitive efficiency for presence judgments than for absence judgments, consistent with previous findings showing lower metacognitive performance for absence (Mazor et al., 2020; Mazor & Fleming, 2020).

Finally, although our model was fitted only to amodal detection decisions, it successfully reproduced modality-specific detection performance, and therefore participants' ability to correctly monitor the source of their percept. These results are in line with a recent study showing that information from different sensory modalities is processed separately before being integrated to reach a common decision threshold (Egan et al., 2025). It also reproduced modality-specific confidence ratings for the two modalities, with confidence defined as the probability of being correct at the time of the decision. This result corroborates previous findings showing that, in some settings, confidence closely matches the probability of being correct, especially when decision and confidence are reported simultaneously (Pouget et al., 2016; Aitchison et al., 2015).

In summary, during an audiovisual detection task at unimodal near-threshold intensity, although the presence of two sensory sources of evidence instead of one improved detection performance, it did not improve metacognitive efficiency. Our ideal observer model, equipped with two modality-specific sensors and making decisions based on a disjunctive integration rule, successfully reproduced amodal and modality-specific detection and modality-specific confidence ratings. However, amodal confidence ratings were successfully reproduced only when presence and absence judgments relied on distinct integration rules. This shows that intersensory congruency plays a critical role in confidence judgments. Overall, these findings indicate that different integration rules apply to perceptual and metacognitive decisions and underscore the importance of counterfactual reasoning for absence judgments. They further suggest that – counterintuitively – a multisensory context might not only impact the perception that something is present, but also the perception that something is absent.

Method

Protocol

In two pre-registered online experiments (**Exp.1**: *https://osf.io/3nvyx*, **Exp.2**: *https://osf.io/ehndv*), participants performed an audiovisual detection task in which they indicated whether a stimulus was present or absent regardless of the modality of presentation. Data were collected from participants recruited via the Prolific platform (N = 60 in Experiment 1; N = 61 in Experiment 2). Participants listened to auditory pink noise while observing dynamic visual Gaussian noise, updated every 33 ms. In Experiment 1, a visual stimulus (a light gray circle spanning 7.5% of the screen size and presented at the center of the screen) was embedded in visual noise on half of the trials. Independently, an auditory stimulus (a sinusoidal tone of 1 kHz) was embedded in auditory noise on half of the trials. As a result, a signal was present on 75% of the trials, with 25% of the trials including both a visual and an auditory signal presented simultaneously (AV trials), and 25% of the trials including none (absent trials). In Experiment 2, we increased the proportion of trials without any signal to obtain an equal proportion of target absent and target present trials. As a result, half of the trials contained no stimulus (absent trials), and the other half contained a stimulus in either the auditory, visual or audiovisual modality in equal proportion (present trials).

When launching the experiment, participants had to calibrate the volume of the auditory noise to a comfortable level. Then, to check if they were using headphones instead of speakers, they had to judge if a sound was presented to their left or right ear in three trials. The experiment was aborted if one error was made.

Before starting the main task, stimulus intensity was calibrated for each participant to reach a 50% detection rate in visual and auditory conditions, based on unimodal psychometric curves. To compute these psychometric curves, participants had to detect stimuli of five different intensities, each presented 20 times, with an additional 20 trials where no stimulus was presented. Auditory and visual calibrations were conducted separately, with the order of presentation counterbalanced across participants.

Each calibration could be repeated if the detection threshold could not be estimated by the fitting procedure, a sign of poor behavior. Pilot study revealed an increase in stimulus detectability between the psychometric evaluation and the main experimental phase. To account for this increase and ensure a similar number of hit and miss trials during the main experiment, we defined stimulus intensity for the main experiment as the intensity corresponding to 40% of detected stimuli during the psychometric evaluation (see SI for individual psychometric curves).

Following ten training trials, participants undertook the main task. On each trial, the stimulus could appear 200, 300, or 400ms after the noise onset and was presented for 600ms. Participants could respond from the start of the trial until a 4-s time limit after the stimulus offset. They first pressed the right or left arrow on the keyboard to indicate whether a stimulus was present or absent irrespective of its sensory modality (the response key to report the presence or absence of the stimulus was counterbalanced across participants). Critically, they were instructed to report presence if a stimulus was present visually, auditorily, or both. Following a 100 ms delay, they reported their amodal confidence in their detection choice by moving a cursor with the mouse on a scale from 0 (“sure incorrect”) to 100 (“sure correct”). Participants were instructed to report an amodal confidence judgement reflecting decision accuracy irrespective of the sensory modality. Finally, they were asked to report their modality-specific detection and confidence judgments on a bi-dimensional (audio/visual) scale, with each axis ranging from 100% sure not perceived to 100% sure perceived, and corresponding to one modality. The mapping of auditory and visual modalities to the horizontal and vertical axes was counterbalanced across participants, although this counterbalancing was not preregistered for Experiment 1.

Participants performed 288 trials in Experiment 1 and 252 in Experiment 2, divided into six experimental blocks, each experimental condition was randomly presented within each block.

In Experiment 1, six participants were excluded based on our pre-registered exclusion criteria (no variability in confidence judgments and no convergence of psychometric curves). An additional five participants were excluded as their detection accuracy was 0% in either the auditory or the visual modality. As trials with confidence ratings below 50 were not analyzed, one participant was excluded as they responded with a confidence below 50 in 286 out of 288 trials. In Experiment 2, three participants were excluded based on our pre-registered exclusion criteria and an additional four participants were excluded as their detection accuracy was 0% in either the auditory or the visual modality. As a result, 48 participants were included in the main analysis for Experiment 1, and 54 for Experiment 2.

Data analysis

All statistical analyses were performed in R. To investigate amodal detection performance and metacognitive sensitivity, we conducted a mixed-effects logistic regression on accuracy as a function of experimental conditions (stimulus absent, auditory, visual, or audiovisual), amodal confidence, and their interaction. To compare reaction times, we conducted a mixed-effects linear regression of log-transformed RTs as a function of the experimental condition and type of judgment (results in SI).

To investigate modality-specific detection and confidence effects, we pre-registered four models depending on whether the stimulus was present or absent in the given modality, and whether it was judged as absent or present in that modality. For consistency with our amodal analyses, we deviated from the original pre-registration and applied the same model structure as for the amodal effects, that is a mixed-effects logistic regression on visual accuracy as a function of experimental condition, visual confidence, and their interaction, and a mixed-effects logistic regression on auditory accuracy as a function of experimental condition, auditory confidence, and their interaction. Full details of the pre-registered analyses are available in the SI.

To investigate confidence bias, we fitted two linear models: one comparing the mean of the confidence ratings as a function of the condition of presentation and response accuracy, and the other comparing the mean of confidence judgments as a function of participants' judgments (stimulus judged absent or present) and accuracy.

To assess participants' ability to monitor their absence judgments, we performed a mixed-effects logistic regression on the accuracy of absence judgments (miss vs correct rejections) as a function of amodal confidence.

Finally, we estimated metacognitive efficiency using response-conditional hierarchical M-ratio (Hmeta-d', Fleming, 2017), both for amodal and modality-specific judgments.

We additionally performed unpreregistered analyses: a Bayesian one-sample t-test on participants' response bias and a Wilcoxon signed-rank test to investigate participants' source monitoring.

Model

The original model (Mazor et al., 2025) considered that when a stimulus is present it remains on the screen until participants made a decision. In contrast, in our experiments, the stimulus was presented for a fixed duration of 600 ms, with temporal uncertainty in its onset. To account for this limited and variable presentation period, we divided the evidence accumulation process into three distinct temporal phases: (1) before the earliest possible stimulus onset, (2) during the window of potential stimulus presentation, and (3) after the stimulus could no longer be present. Before the earliest onset, posterior probabilities updates were driven solely by prior beliefs about presence in each modality. During the stimulus presentation window, updating was based on both prior beliefs and incoming sensory input. After this window, the probability of accumulating evidence for presence decayed. We compared this model to a model without decay, in which evidence for presence was based on random noise during periods when no stimulus could be present, corresponding to the probability of sampling a 1 if the target is absent. We found that the model with a decay mechanism fitted the data better than the model without (difference in AIC: 2593.8).

Model comparison

We compared the goodness of fit to amodal decisions across different combinations of parameters representing the believed and true likelihoods (see Table 2). Model comparisons at the group level showed that the best model was the "single belief" model, according to which the believed probability of sampling a 1 (if a target is present or absent) is the same for auditory and visual signals. The model selection was based solely on the data from Experiment 1, and we tested the best model's ability to generalize to new data with results from Experiment 2.

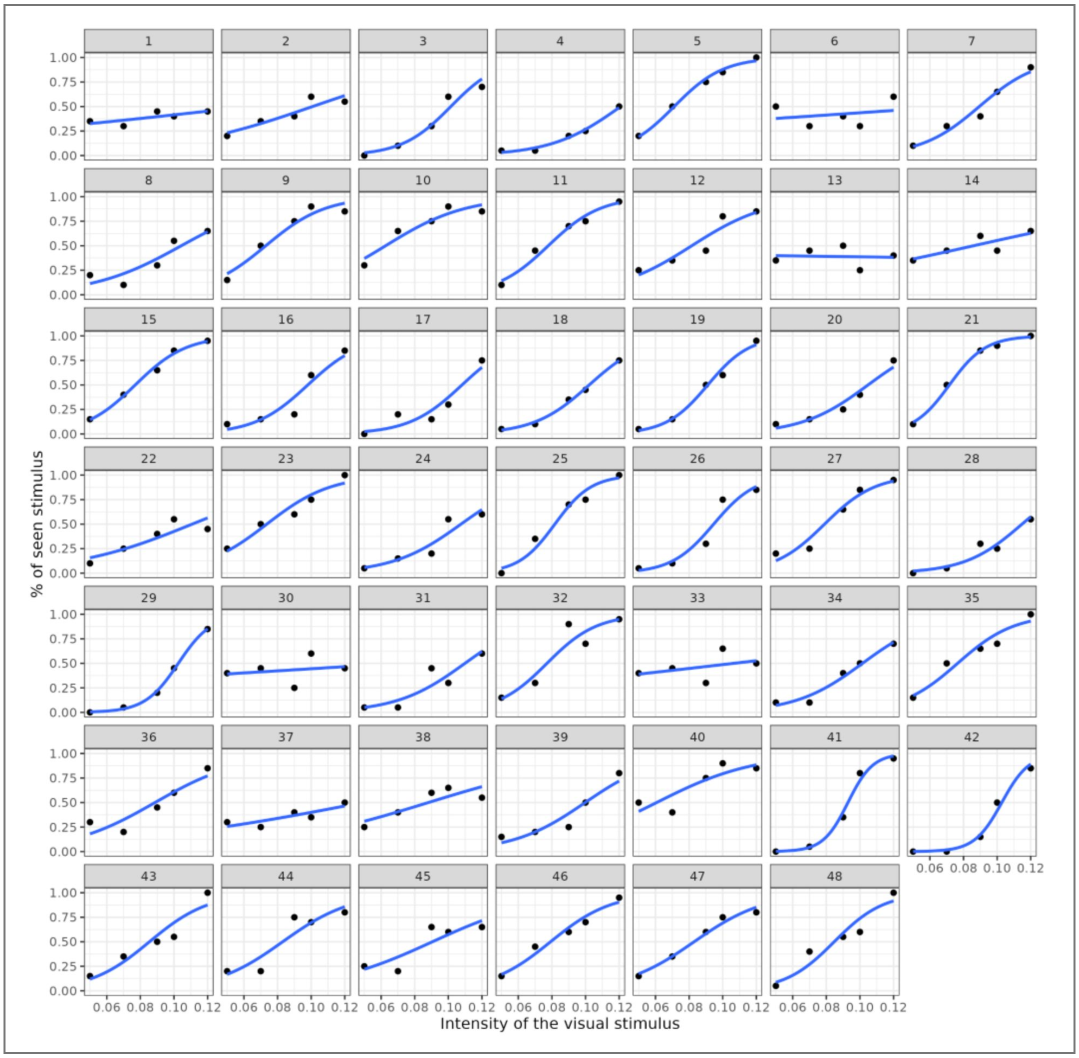
Table 2. Fitted models (rows) with different or identical parameters (columns) across the visual and auditory sensors.

	True thetas $\boldsymbol{\theta}$	Believed thetas $\bar{\theta}$	AIC (relative to the best model)
Single belief	Different	Same	0
Full	Different	Different	22.13
No belief	Different	/	100.94
Single likelihood	Same	Different	161.23
Single belief & likelihood	Same	Same	334.05
No belief & single likelihood	Same	/	336.54

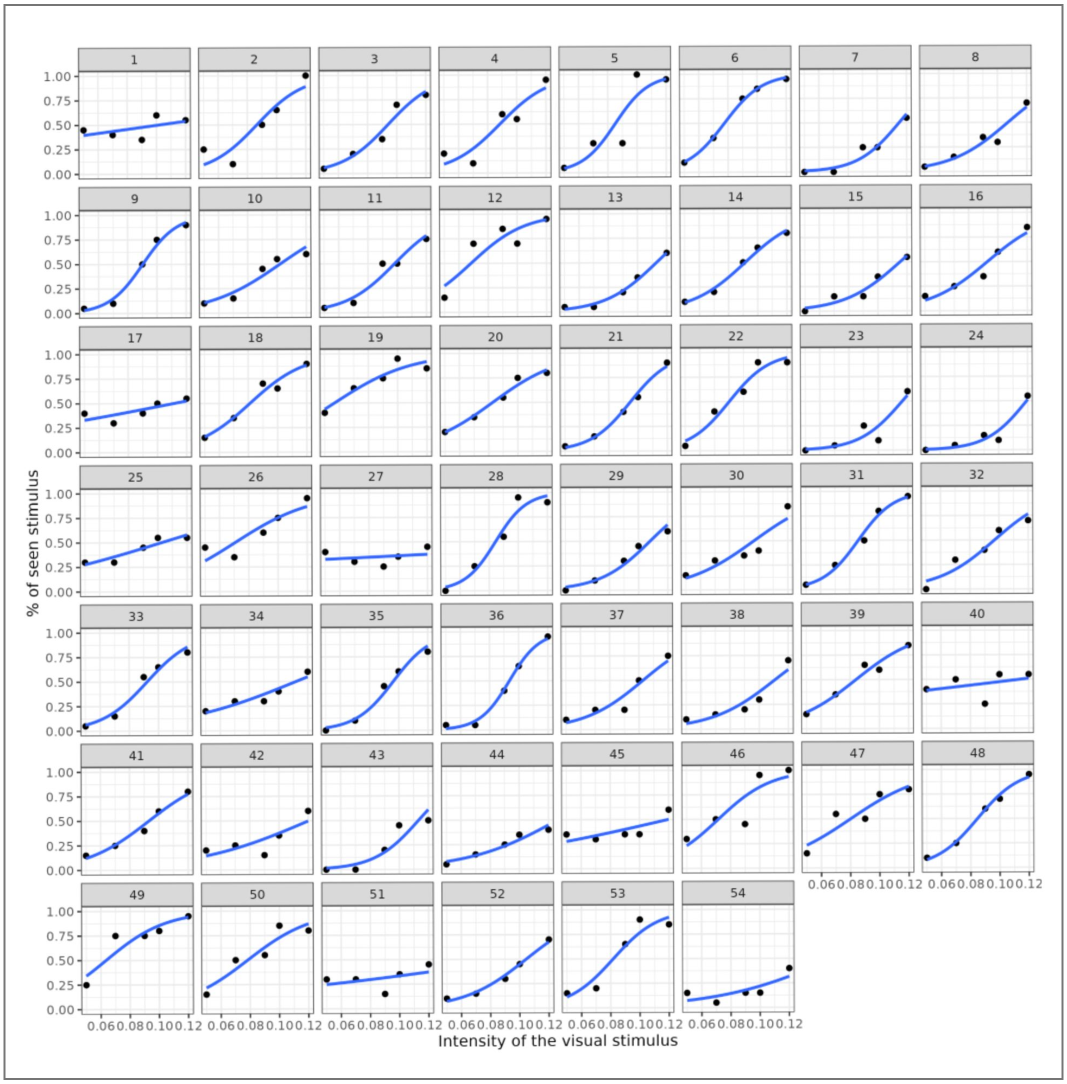
Supplementary information

Psychometric curves

Visual psychometric curves

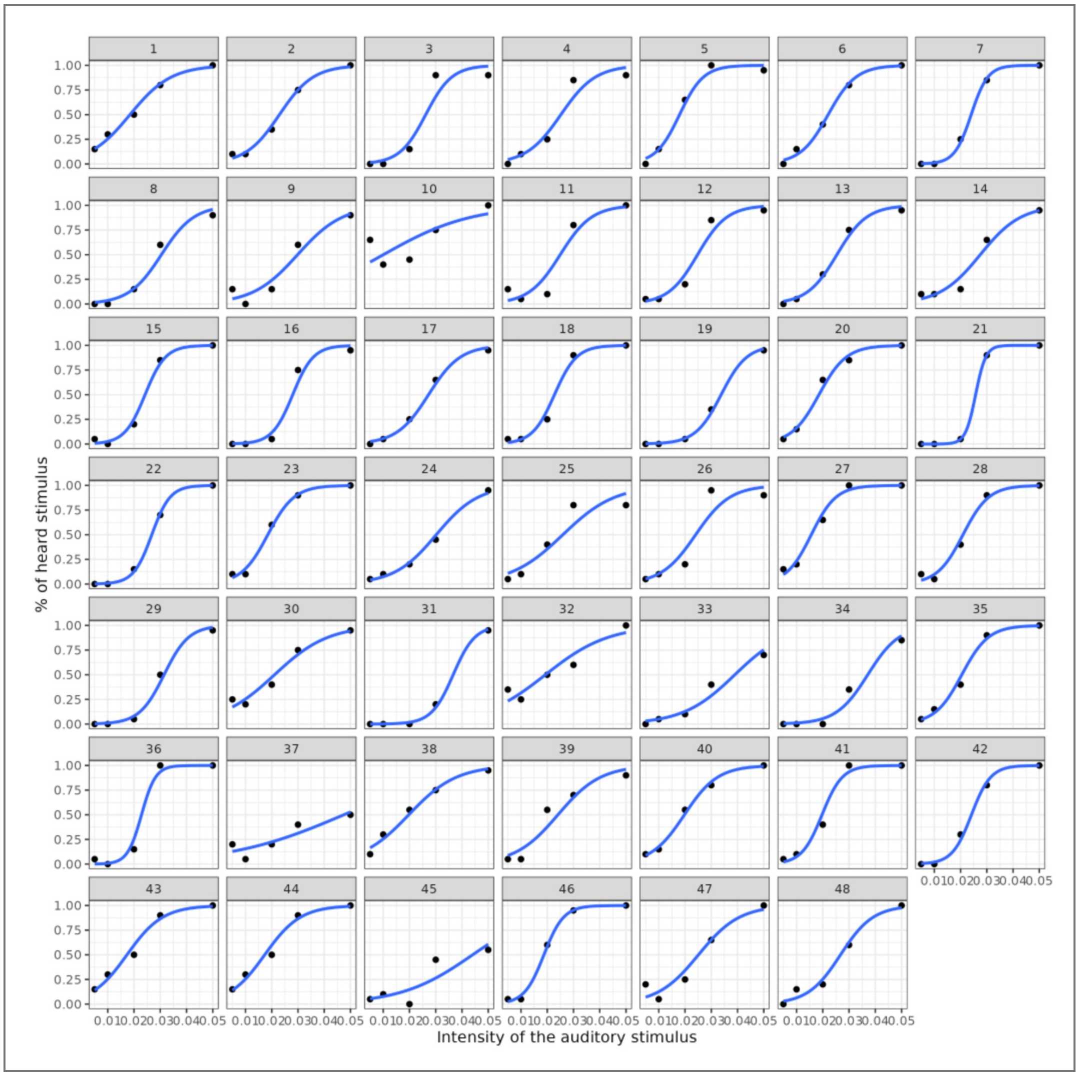


Supplementary figure 1. Visual psychometric curve for each participant of Experiment 1

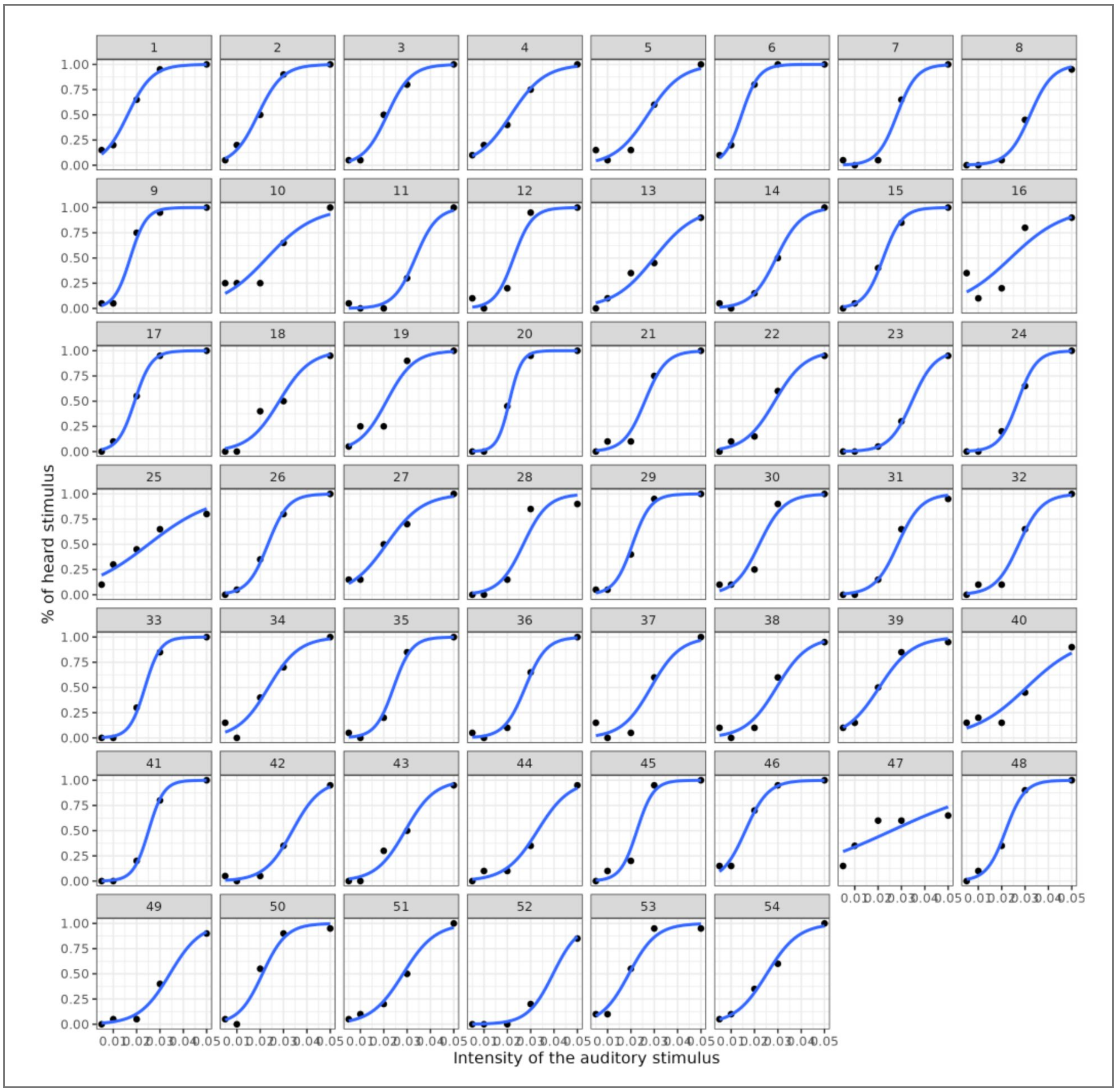


Supplementary figure 2. **Visual psychometric curve for each participant of Experiment 2**

Auditory psychometric curves



Supplementary figure 3. **Auditory psychometric curve for each participant of Experiment 1**



Supplementary figure 4. **Auditory psychometric curve for each participant of Experiment 2**

Data analysis

Contrast

For the model visual accuracy ~ condition * visual confidence, we used the following contrast coding for the condition of presentation: (-1, -1, 1, 1) to compare visually present versus absent trials; (-1, 1, 0, 0) to assess the effect of auditory information on visually absent trials; (0, 0, -1, 1) to assess the effect of auditory information on visually present trials.

For the model auditory accuracy ~ condition * auditory confidence, we used the following contrast coding: (-1, 1, -1, 1) to compare auditory present versus absent trials; (-1, 0, 1,0) to test the influence of visual information on auditory-absent trials; (0, -1, 0, 1) to test the influence of visual information on auditory-present trials.

Additional results

Analysis of Reaction times

In Experiment 1, reaction times did not differ between bimodal and unimodal trials ($\hat{\beta}$ = 0[−0.01,0.02], BF0 = 4.94), between auditory and visual trials ($\hat{\beta}$ = 0.01[−0.02,0.03], BF0 = 4.33), nor between present and absent trials ($\hat{\beta}$ = 0 [0,0], BF0 = 7.17). Evidence for a difference between presence and absence judgments was inconclusive ($\hat{\beta}$ =−0.03 [−0.06,0.01], BF0 = 1.81). However, the difference in response times between absence and presence judgments was larger in bimodal compared to unimodal trials, reflecting a stronger accuracy effect for audiovisual stimuli ($\hat{\beta}$ =−0.02 [−0.03,−0.01], BF1 > 1000). This interaction effect was inconclusive when comparing visual and auditory trials ($\hat{\beta}$ =−0.03 [−0.06,0.01], BF1 = 0.50).

In Experiment 2, evidence was inconclusive for a difference between bimodal and unimodal trials ($\hat{\beta}$ = 0 [0,0], BF0 = 1.48) and for a difference between auditory and visual trials ($\hat{\beta}$ = 0 [−0.01,0.01], BF0 = 1.71). However, participants were faster on present than on absent trials ($\hat{\beta}$ = 0.00 [−0.01,0.00], BF1 = 4.65). Participants were faster for presence compared to absence judgments ($\hat{\beta}$ =−0.05, 95% CI [−0.08, −0.02], BF1 = 125). Moreover, the difference in response times between absence and presence judgments was larger in bimodal compared to unimodal trials, reflecting a stronger accuracy effect for audiovisual stimuli ($\hat{\beta}$ =−0.02, 95% CI [−0.03, −0.02], BF1 > 1000). This effect was also stronger in visual than auditory trials ($\hat{\beta}$ =−0. 04, 95% CI [−0. 06, −0. 02], BF1 = 591).

Finally, in both experiments, we observed a significant interaction between the condition of presentation and the type of judgment, the difference in response times between absence and presence judgments was larger in present compared to absent trials indicating that participants responded faster on correct-present trials (**Exp.1**: $\hat{\beta}$ =−0.03, 95% CI [−0. 04, −0. 02], BF1 > 1000; **Exp.2**: $\hat{\beta}$ =−0. 03, 95% CI [−0. 04, −0. 02], BF1 > 1000).

Source monitoring: Test for a visual dominance effect

A Wilcoxon test showed no difference between the proportion of visual-only and auditory-only reports relative to their predicted probabilities assuming independent visual and auditory detection (**Exp.1:** V = 421, p = .09; **Exp.2:** V = 870, p = .27). This shows that visual categorizations occurred at the level expected from the higher visual hit rate, with no evidence for additional visual dominance.

Response bias toward absence at the modality-specific level

Although participants reported whether they detected a stimulus regardless of modality, the modality-specific scale enabled us to infer detection performance for each modality. As for amodal detection, participants had a tendency to respond that both auditory and visual stimuli were absent, with Bayesian logistic regressions on modality-specific accuracy showing that participants had a higher auditory accuracy for auditory absent compared to auditory present trials (**Exp.1**: $\hat{\beta}$

=−1.69 [−2.07,−1.33], *BF1* > 1000; **Exp.2:** $\hat{\beta}$ =−1.65 [−1.95,−1.34], BF1 > 1000), and higher visual accuracy for visually absent compared to visually present trials (**Exp.1**: $\hat{\beta}$ =−1.75 [−2.19,−1.33], BF1 > 1000; **Exp.2**: $\hat{\beta}$ =−1. 32 [−1. 58, −1. 07], BF1 > 1000).

Multisensory interference

Influence of the visual modality on auditory judgments

In Experiment 1, participants were more accurate at detecting the absence of the auditory stimulus when the visual stimulus was also absent ($\hat{\beta}$ =−0.21 [−0.32, −0.1], BF1 = 12.14). Evidence was inconclusive in Experiment 2 ($\hat{\beta}$ =−0.13 [−0.25,−0.03], BF1 = 0.70). In both experiments, detecting the presence of the auditory stimulus was not significantly impacted by the presence or absence of a visual stimulus (**Exp.1**: $\hat{\beta}$ = 0.09 [0.02,0.16], BF1 = 1.85; **Exp.2**: $\hat{\beta}$ = 0.06 [−0.01,0.14], BF1 = 0.80) (see Supp. Fig.5A). These suggest that the presence of a visual stimulus may bias participants toward responding that something is present at the auditory level only when the auditory stimulus is actually absent.

At the metacognitive level, results suggest a facilitative effect of audiovisual congruency. Auditory metacognitive sensitivity for absence was higher when there was also no visual stimulus (**Exp.1**: $\hat{\beta}$ =−0.02 [−0.03,0], BF1 = 3.19; **Exp.2**: $\hat{\beta}$ =−0.01 [−0.02,0.00], BF1 = 15.39). Results for auditory metacognitive sensitivity for presence are more mixed with Experiment 2 showing a higher auditory metacognitive sensitivity when there was also a visual stimulus ($\hat{\beta}$ = 0.01 [0.01,0.02], BF1 = 20.54), while evidence from Evidence 1 was inconclusive ($\hat{\beta}$ = 0.01 [0,0.02], BF1 = 0.53) (see Supp. Fig.5B).

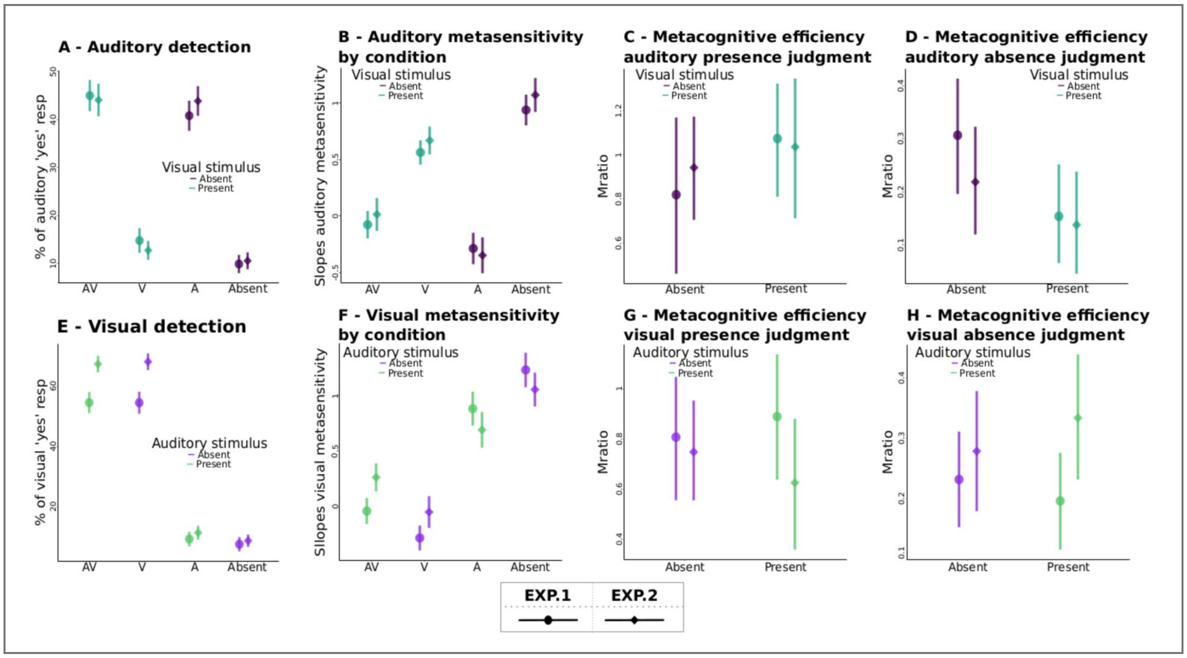
We further assessed auditory metacognitive efficiency as a function of the visual modality. In both experiments, when the auditory stimulus was judged present, the presence of a visual stimulus had no effect on metacognitive efficiency (**Exp.1**: Δ*M* = 0.24 [-0.19, 0.68]; **Exp.2**: Δ*M* = 0.10 [-0.30, 0.48]). In contrast, in Experiment 1, when the auditory stimulus was judged absent, metacognitive efficiency was significantly higher when the visual stimulus was absent (Δ*M* = -0.16 [-0.30, -0.01]). However, this effect did not reach significance in Experiment 2 (*M* = Δ-0.08 [-0.21, 0.06]) (see Supp. Fig.5C-D).

Influence of the auditory modality on visual judgments

We tested if the auditory modality also biased participants toward responding that something is present at the visual level only when the visual stimulus was actually absent. While this effect was inconclusive in Experiment 1 ($\hat{\beta}$ =−0.17 [−0.32,−0.02], *BF1* = 1.18), participants were more accurate at detecting the absence of the visual stimulus when there was also no auditory stimulus in Experiment 2 ($\hat{\beta}$ =−0.16 [−0.28, −0.05], BF1 = 6.67). In contrast, in both experiments, detecting the presence of the visual stimulus was not significantly impacted by the presence or absence of the auditory stimulus (see Supp. Fig. 5E).

Visual metacognitive sensitivity for absence was higher when there was also no auditory stimulus in Experiment 2 ($\hat{\beta}$ =−0.02 [−0.03,0.00], BF1 = 6.05), while the evidence was inconclusive in Experiment 1 ($\hat{\beta}$ =−0.01 [−0.03,0], BF1 = 1.64). In contrast, in both experiments, visual metacognitive sensitivity for presence was higher when an auditory stimulus was present (**Exp.1**: $\hat{\beta}$= 0.01 [0.00,0.02], BF1 = 4.52; **Exp.2**: $\hat{\beta}$= 0.01 [0.01,0.02], BF1 = 42.27). These findings mirror the audiovisual congruency effects observed at the auditory level, suggesting cross-modal facilitation when both modalities provide consistent cues (see Supp. Fig. 5F).

We also examined visual metacognitive efficiency as a function of the auditory modality. In both experiments, there was no clear effect of the auditory modality when the visual stimulus was judged absent (**Exp.1**: Δ*M* = -0.04 [-0.15, 0.08]; **Exp.2:** Δ*M* = 0.06 [-0.09, 0.20]), nor when it was judged present (**Exp.1**: Δ*M* = 0.09 [-0.26, 0.44]; **Exp.2**: Δ*M* = -0.13 [-0.46, 0.20]) (see Supp. Fig. 5G-H).



Supplementary figure 5. **Modality-specific results.**

Experiment 1 is represented by circles, while Experiment 2 is represented by diamonds. A) Auditory detection performance: Percentage of stimuli judged to be present at the auditory level as a function of the experimental condition; error bars represent the standard error. B) Auditory metacognitive sensitivity as a function of the experimental condition; error bars represent the standard error. C) Auditory metacognitive efficiency (response-conditional Mratio) for auditory present judgments as a function of the visual modality; error bars represent the highest density interval. D) Auditory metacognitive efficiency (response-conditional Mratio) for auditory absent judgments as a function of the visual modality; error bars represent the highest density interval. E) Visual detection performance: Percentage of stimuli judged to be present at the visual level as a function of the experimental condition; error bars represent the standard error. F) Visual metacognitive sensitivity as a function of the experimental condition; error bars represent the standard error. G) Visual metacognitive efficiency (response-conditional Mratio) for visual present judgments as a function of the auditory modality; error bars represent the highest density interval. H) Visual metacognitive efficiency (response-conditional Mratio) for visual absent judgments as a function of the auditory modality; error bars represent the highest density interval.

Metacognitive sensitivity between presence and absence

We preregistered the estimation of metacognitive sensitivity conditional to the stimulation conditions. We found that amodal metacognitive sensitivity was higher for absent than for present trials, indicating more accurate confidence calibration in trials where no stimulus was present (**Exp.1:** $\hat{\beta}$ =−0.27 [−0.4, −0.13], *BF1* = 245; **Exp.2:** $\hat{\beta}$ =−0. 23[−0.34,−0.12], BF1 > 1000).

We observed a similar pattern at the modality-specific level. Participants had a higher auditory metacognitive sensitivity for auditory absent than for auditory present trials (**Exp.1**: $\hat{\beta}$ =−0. 04 [−0.07,−0.02], BF1 = 130; **Exp.2**: $\hat{\beta}$ =−0. 05 [−0.06,−0.03], BF1 = 270), and a higher visual metacognitive sensitivity for visually absent compared to visually present trials (**Exp.1**: $\hat{\beta}$ =−0.05 [−0.08,−0.03], BF1 > 1000; **Exp.2**: $\hat{\beta}$ =−0.04 [−0.06,−0.02], BF1 > 1000). Additionally, participants demonstrated metacognitive insight into their amodal absence responses: they were more confident in their correct rejections than in their misses (**Exp.1**: $\hat{\beta}$ =−0.41 [−0.51,−0.31], BF1 > 1000; **Exp.2**: $\hat{\beta}$ =−0.39 [−0.49, −0.3], BF1 > 1000).

We additionally investigated response-conditional metacognitive sensitivity to investigate differences between presence and absence judgments. Response-conditional amodal metacognitive sensitivity showed no difference between presence and absence judgements in Experiment 1 ($\hat{\beta}$ = 0.01 [−0.21,0.21], BF0 = 6.44), and higher metacognitive sensitivity for presence than for absence judgments in Experiment 2 ($\hat{\beta}$= 0.27 [0.11,0.43], BF1 = 57.58). At the modality-specific level, response-conditional results were inconclusive for a difference between presence and absence judgments both at the auditory (**Exp.1:** $\hat{\beta}$=0 [0,0.1], BF1 = 1.32; **Exp.2**: $\hat{\beta}$=0.01 [−0.01,0.02], BF1 = 1.04) and at the visual level (**Exp.1:** $\hat{\beta}$ = 0 [−0.01,0.01], BF1 = 0.97; **Exp.2**: $\hat{\beta}$ = 0 [−0.01,0.01], BF1 = 0.92).

Preregistered modality-specific analysis

Auditory modality

In experiment 1, when the audio stimulus was present, participants were better at detecting it as being auditory present when a visual stimulus was present ($\hat{\beta}$ = 0.20, 95% CI [0.08,0.32], z = 3.24, p=.001). Moreover, participants adapted better their auditory confidence to their auditory accuracy for audiovisual trials compared to auditory trials ($\hat{\beta}$ = 0.02, 95% CI [0.00,0.04], z = 2.39, *p* =.017). When the audio stimulus was absent, the presence of a visual stimulus biased participants toward judging the audio stimulus as being present ($\hat{\beta}$ = 0.41, 95% CI [0.22,0.59], *z* = 4.30, *p* <.001). Moreover, participants adapted better their auditory confidence to their auditory accuracy for fully absent trials compared to trials in which the audio stimulus was absent but the visual stimulus was present ($\hat{\beta}$ = 0.03, 95% CI [0.01,0.04], *z* = 3.41, *p* <. 001).

When the audio stimulus was judged present, participants were able to distinguish between their hits and the false alarms ($\hat{\beta}$ = 0.05, 95% CI [0.03,0.06], *z* = 7.43, *p* <. 001). When the audio stimulus was judged absent, participants were able to distinguish between their correct rejections and their misses ($\hat{\beta}$ =−0.02, 95% CI [−0.03,−0.01], *z* =−4.83, *p* <. 001). However, this difference decreased in the presence of a visual stimulus ($\hat{\beta}$ = 0.02 , 95% CI [0. 01, 0. 03], *z* = 5. 36, p <. 001).

In experiment 2, when the audio stimulus was present, participants adapted better their auditory confidence to their auditory accuracy for audiovisual compared to audio trials ($\hat{\beta}$ = 0.04, 95% CI [0.01,0.06], *z* = 2.91, *p* =.004). There was no main effect of visual presence on detection. When the audio stimulus was absent, participants adapted better their auditory confidence to their auditory accuracy for fully absent trials compared to trials in which the audio stimulus was absent but the visual stimulus was present ($\hat{\beta}$ = 0.02, 95% CI [0.00,0.05], *z* = 2.18, *p* =. 029). There was no main effect of visual presence on absence detection.

When the audio stimulus was judged present, participants were able to distinguish between their hits and false alarms at the auditory level ($\hat{\beta}$ = 0.05, 95% CI [0.04,0.07], *z* = 6.74, *p* <. 001). When the audio stimulus was judged absent, participants were able to distinguish between their correct rejection and their misses at the auditory level ($\hat{\beta}$ =−0.02, 95% CI [−0.03,−0.01], z =−5.07, p <.001). However, this difference decreased in the presence of a visual stimulus ($\hat{\beta}$ = 0.01, 95% CI [0.00,0.02], z = 2.61, p =. 009).

Visual modality

In experiment 1, when the visual stimulus was present, participants adapted better their visual confidence to their visual accuracy for audiovisual trials compared to visual ones ($\hat{\beta}$ = 0.02, 95% CI [0.01,0.03], z = 2.73, p =.006). There was no main effect of auditory presence on visual detection. When the visual stimulus was absent, the presence of an auditory stimulus biased participants toward judging the visual stimulus as being present ($\hat{\beta}$ = 0.36, 95% CI [0.11,0.61], z = 2.80, p =.005). Moreover, participants adapted better their visual confidence to their visual accuracy for fully absent trials compared to trials in which the visual stimulus was absent but the auditory stimulus was present ($\hat{\beta}$ = 0.03, 95% CI [0. 01, 0. 05], z = 2. 97, p =. 003).

When the visual stimulus was judged present, participants were able to distinguish between their hit and the false alarms ($\hat{\beta}$ = 0.06, 95% CI [0.05,0.08], z = 7.43, p <. 001). When the visual stimulus was judged absent, participants were able to distinguish between their correct rejection and their misses ($\hat{\beta}$ =−0.03, 95% CI [−0.04,−0.02], z =−7.07, p <.001). However, this difference decreased in the presence of an auditory stimulus ($\hat{\beta}$ = 0. 01, 95% CI [0. 00, 0. 02], z = 2. 26, p =. 024).

In experiment 2, when the visual stimulus was present, participants adapted better their visual confidence to their visual accuracy for audiovisual trials compared to visual ones ($\hat{\beta}$ = 0.03, 95% CI [0.01,0.05], z = 2.94, p =.003). There was no main effect of auditory presence on detection. When the visual stimulus was absent, the presence of an auditory stimulus biased participants toward judging the visual stimulus as being present ($\hat{\beta}$ = 0.31, 95% CI [0.09,0.54], z = 2.80, p =.005). Moreover, participants adapted better their visual confidence to their visual accuracy for fully absent trials compared to trials in which the visual stimulus was absent but the auditory stimulus was present ($\hat{\beta}$ = 0.03, 95% CI [0. 00, 0. 06], z = 2. 22, p =. 026).

When the visual stimulus was judged present, participants were able to distinguish between their hit and the false alarms($\hat{\beta}$ = 0.07, 95% CI [0.05,0.09], z = 7.10, p <. 001). When the visual stimulus was judged absent, participants were able to distinguish between their correct rejection and their misses ($\hat{\beta}$ =−0.04, 95% CI [−0.06,−0.03], z =−6.45, p <. 001). There was no effect of the presence of auditory stimuli on this distinction.

Predicting Amodal Confidence from Modality-Specific Confidence Ratings

We pre-registered the investigation of the contribution of auditory and visual confidence to amodal confidence. Beyond this, we explored which model best predicted amodal confidence from modality-specific confidence. We compared different integration rules:

1. Max model: amodal confidence ~ max(auditory confidence, visual confidence)
2. Linear model: amodal confidence ~ auditory confidence * visual confidence (equal weight for each modality)
3. Weighted linear model: amodal confidence ~ auditory confidence_w * visual confidence_w (the weights of the auditory and visual confidence are determined by the data, thus allowing the relative contribution of each modality to vary independently)
4. Optimal integration model: amodal confidence ~ auditory confidence + visual confidence - auditory confidence*visual confidence
5. Weighted min-max model: max(auditory confidence, visual confidence) + min(auditory confidence, visual confidence)

Based on this first comparison, we found that the min-max model was the one that best explained amodal confidence. We compared different ways to improve this min-max model by adding:

1. An absolute confidence score computed as (absolute(minimal_confidence −0.5)+absolute(maximal_confidence −0.5)) /2: amodal confidence ~ max(auditory confidence, visual confidence) + min(auditory confidence, visual confidence) + absolute confidence

2. Absolute confidence and type of judgments: amodal confidence ~ max(auditory confidence, visual confidence) + min(auditory confidence, visual confidence) + absolute confidence*judgments
3. Optimal integration of the min-max: amodal confidence ~ max(auditory confidence, visual confidence) + min(auditory confidence, visual confidence) - max(auditory confidence, visual confidence)*min(auditory confidence, visual confidence)

Models comparison showed that the best model was the min-max with the absolute confidence and the type of judgments taken into account.

Supplementary table 1. Comparison of the different models tested.

Model tested	AIC (relative to the best model)	BIC (relative to the best model)	Free parameters	Marginal R2
Absolute confidence & type of judgment	0	0	13	*Exp1* : 0.86 *Exp2*: 0.83
Absolute confidence	*Exp1*: 11889.62 *Exp2*: 10876.96	*Exp1* : 11859.70 *Exp2*: 10847.16	9	*Exp1* : 0.71 *Exp2*: 0.74
Min-max following optimal integration	*Exp1*: 11981.87 *Exp2*: 10978.72	*Exp1* : 11951.95 *Exp2*: 10948.92	9	*Exp1* : 0.79 *Exp2*: 0.81
Min-max integration	*Exp1*: 12074.96 *Exp2*: 11050.43	*Exp1* : 12030.08 *Exp2*: 11005.73	7	*Exp1* : 0.76 *Exp2*: 0.81
Max model	*Exp1*: 12074.96 *Exp2*: 11548.43	*Exp1* : 12458.14 *Exp2*: 11488.82	5	*Exp1* : 0.68 *Exp2*: 0.74
Optimal integration	*Exp1*: 14488.18 *Exp2*: 13035.70	*Exp1* : 14428.35 *Exp2*: 12976.09	5	*Exp1* : 0.64 *Exp2*: 0.75
Linear model with variable weight	*Exp1*: 19993.53 *Exp2*: 15516.70	*Exp1* : 19948.65 *Exp2*: 15471.99	7	*Exp1* : 0.71 *Exp2*: 0.81
Linear model with equal weight	*Exp1*: 20847.36 *Exp2*: 16832.97	*Exp1* : 20787.52 *Exp2*: 16773.37	5	*Exp1* : 0.67 *Exp2*: 0.72

Closer examination of this model's parameter estimates revealed a significant main effect of the maximal confidence (**Exp.1**: $\hat{\beta}$ = 0.15, 95% CI [0.10,0.20], t(46.96) = 5.57, p <. 001; **Exp.2**: $\hat{\beta}$ = 0.22, 95% CI [0.17,0.27], t(53.73) = 8.19, p <. 001), of the minimal confidence (**Exp.1**: $\hat{\beta}$= 0.08, 95% CI [0.05,0.11], t(40.15) = 5.17, p<.001; **Exp.2**: $\hat{\beta}$ = 0.08, 95% CI [0.06,0.10], t(38.08) = 7.67, p<.001), and of the absolute of confidence (**Exp.1**: $\hat{\beta}$= 0.07, 95% CI [0.02,0.11], t(43.91) = 2.69, p=.010; **Exp.2**: $\hat{\beta}$ = 0.07, 95% CI [0.03,0.11], t(50.35) = 3.73, p <.001), as well as a strong interaction effect between the type of judgments and the absolute confidence (**Exp.1**: $\hat{\beta}$ = 0.96, 95% CI [0.79,1.12], t(47.90) = 11.42, p <. 001; **Exp.2:** $\hat{\beta}$= 0.82, 95% CI [0.70,0.95], t(55.86) = 12.99, p <. 001). Altogether, these results indicate that modality-specific confidence signals effectively predict amodal confidence.

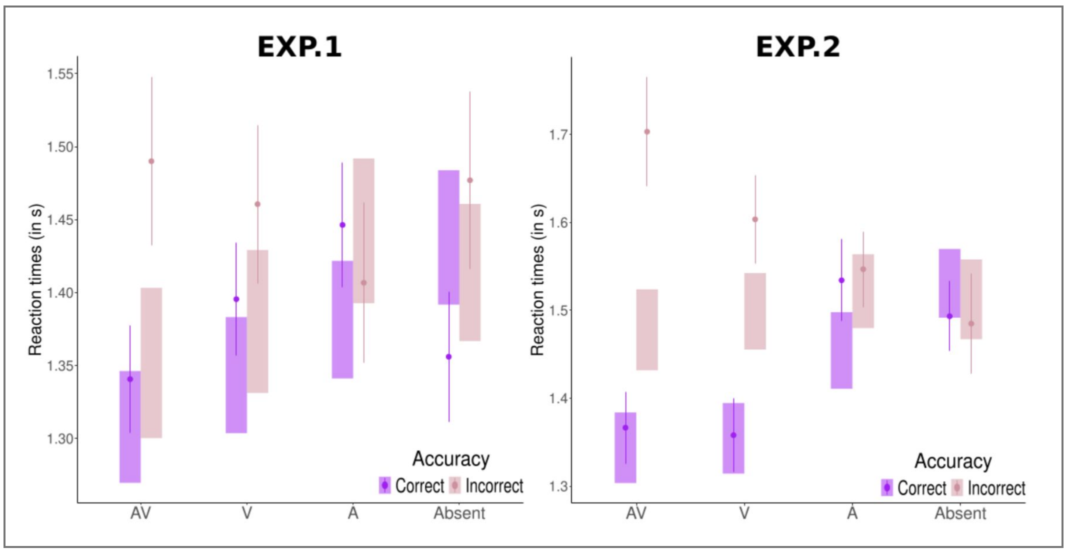
Model

Disjunctive versus conjunctive rule for perceptual detection decision

We compared the goodness of fit to amodal decisions of a model integrating information based on a disjunctive rule to one integrating information based on a conjunctive rule. The best model was the one integrating information based on a disjunctive rule (difference in AIC: 334.37).

Predictions of reaction times

Our model reproduced the observed reaction times only for correct presence judgments.



Supplementary figure 6. **Reproduction of reaction times.**

Reaction times (in seconds) as a function of the condition of presentation for Experiments 1 and 2. Error bars represent the standard error from the data. Rectangles represent data simulated from the model, centered on the mean value and with height equal to the standard error

Parameters recovery

To test the recoverability of our estimated parameters, we repeated the fitting procedure for the simulated data from parameters fitted to our Experiment 1. The correlations between the two sets of parameters were generally high.

Results Experiment 3

The goal of this experiment was to investigate the effect of stimulus intensity on confidence bias towards absence. To examine this, we replicated Experiment 2 while increasing the stimulus intensity to target a higher detection rate of approximately 70% in unimodal trials. All other aspects of the experiment remain the same as in Experiment 2. We preregistered this experiment (*https://osf.io/v5kqd*).

Detection effects

We compared the amodal criterion to 0 to investigate the response bias of participants. We found that participants had a significant bias toward responding that nothing was present (M = 0.21, 95% CI [0.05,0.37], $t(32)$ = 2.70, p =.011). Consequently, they were more accurate for absent compared to present trials ($\hat{\beta}$ =−0.28, 95% CI [−0.46,−0.09], z =−2.91, p =. 004). This was also the case at the modality-specific level where participants had a higher detection rate for auditory absent than for auditory present trials ($\hat{\beta}$ =−1.02, 95% CI [−1.44, −0.60], z =−4.77, p <. 001), and for visually absent than for auditory present trials ($\hat{\beta}$ =−1.20, 95% CI [−1.60, −0.81], z =−5.93, p <. 001).

Looking at multisensory effects, we found that participants detected better bimodal than unimodal trials ($\hat{\beta}$ = 0.59, 95% CI [0.49,0.69], z = 11.36, p <.001). There was no significant difference between visual and auditory trials.

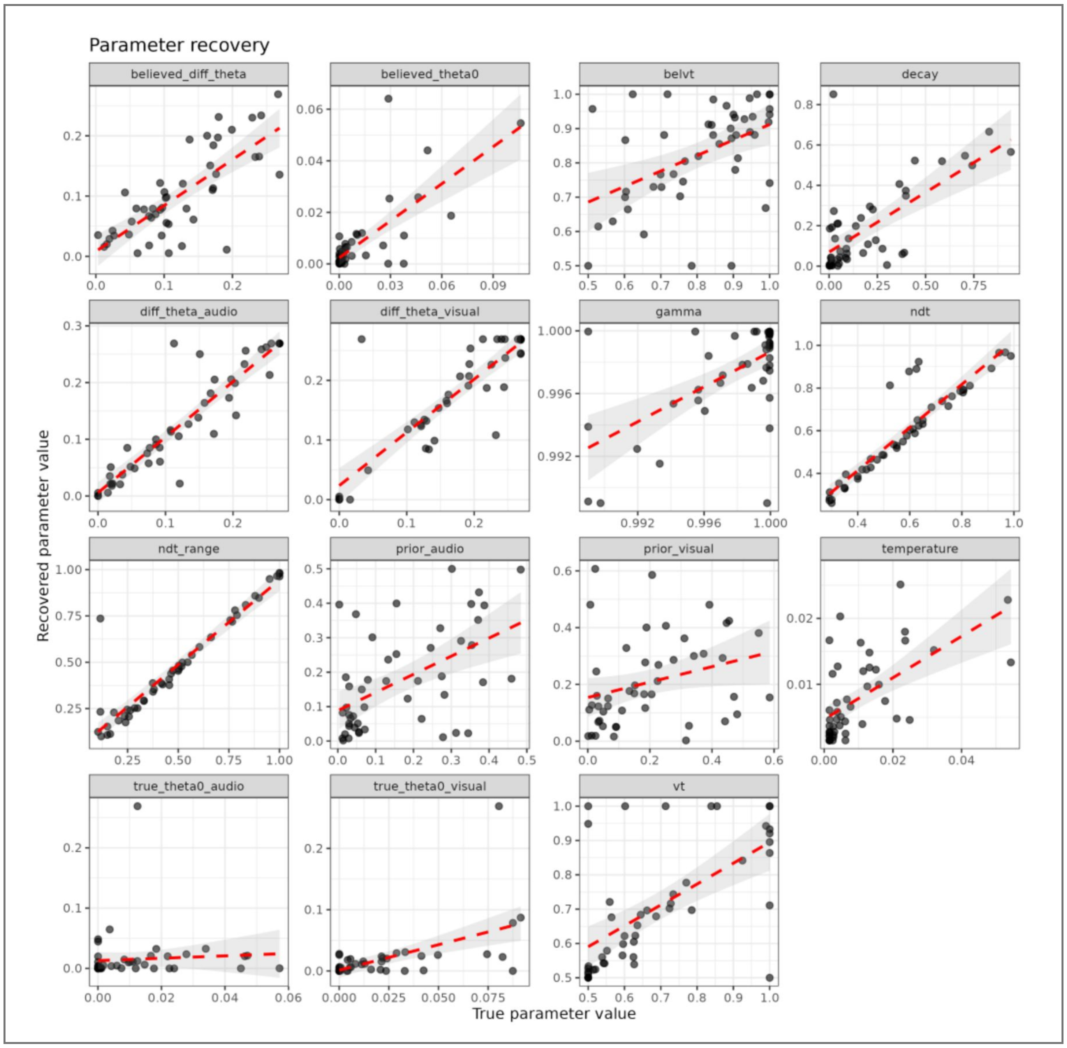
Finally, we compared the reaction times (in seconds), as a function of the experimental condition and as a function of the judgments of participants. We found no significant difference as a function of the condition of presentation. However, participants were faster following a presence judgments than an absence judgments ($\hat{\beta}$ =−0.11, 95% CI [−0.16,−0.07], t(35.65)=−4.82, p <.001). The interaction between the condition and the judgment was significant indicating that participants were faster for present trials compared to absent trials when they judged the stimulus to be present (so when their answer was correct; $\hat{\beta}$ =−0.05, 95% CI [−0.06, −0.04], t(33.03) =−6.96, p <. 001).

Amodal confidence effects

To investigate confidence bias, we compare the mean of the confidence judgments in function of the condition and of the accuracy of the response. We found no main effect of the condition of presentation. Participants were more confident in the correct responses ($\hat{\beta}$= 5.34, 95% CI [3.80,6.88], t(200.50) = 6.80, p <.001). The interaction between the condition and the accuracy of the answer was significant. Planned contrast showed that the effect of accuracy was higher for absent compared to present trials ($\hat{\beta}$ =−1.46, 95% CI [−2.32,−0.61], t(200.34)=−3.36, p <.001), with no difference as a function of the modality of presentation.

We also compare the mean of confidence judgments as a function of the participant's judgment and of the accuracy. We found that participants were more confident for absence compared to presence judgments ($\hat{\beta}$ =−3.07, 95% CI [−4.98, −1.16], t(94.15) =−3.15, p =.002), and more confident when their responses were correct ($\hat{\beta}$ = 6.60, 95% CI [4.69,8.51], t(94.15) = 6.77, p <.001). The interaction between the two was significant and showed a greater effect of the accuracy when the stimulus is judged present compared to judged absent ($\hat{\beta}$ = 4.44, 95% CI [0.62,8.26], t(94.15) = 2.28, p =. 025). We looked at whether participants were able to distinguish between their correct and incorrect absence judgments. We found that participants were more confident in their correct rejections than in their misses ($\hat{\beta}$ =−0.47, 95% CI [−0.63, −0.30], z =−5. 56, p <. 001)

We additionally investigated metacognitive sensitivity to evaluate how accurately participants' confidence tracked their accuracy. We found that participants adapted better their confidence to the accuracy of their answer for absent than for present trials ($\hat{\beta}$ =−0.15 , 95% CI [−0.26,−0.03], z



Supplementary figure 7. **Parameter recovery.**

=−2.50, p =.012), and for bimodal compared to unimodal trials ($\hat{\beta}$ = 0.16, 95% CI [0.06,0.25], z = 3.24, p =. 001). There was no significant difference between visual and auditory trials.

Modality-specific confidence effects

At the auditory level, participants adapted better their auditory confidence to their auditory accuracy for absent than for present trials ($\hat{\beta}$ =−0.05, 95% CI [−0.08, −0.03], z =−3.67 , p <. 001). We also found that participants were more confident in their absence than in their presence judgments ($\hat{\beta}$ =−4.64, 95% CI [−6.43,−2.85], t(210.76) =−5.08, p <.001). There was a main effect of the presence of a visual stimulus, indicating that participants were more confident in the absence of a visual stimulus ($\hat{\beta}$ =−2.78, 95% CI [−4.56,−1.00], t(210.32) =−3.06, p =. 003). However, the interaction between the presence of visual stimulus and the type of judgments was not significant.

At the visual level, participants adapted better their visual confidence to their visual accuracy for absent than for present trials ($\hat{\beta}$ =−0.05, 95% CI [−0.08, −0.03], z =−4.49 , p <.001). In absence, they adapted it better when there was no auditory stimulus ($\hat{\beta}$ =−0.02, 95% CI [−0.04,0.00], z =−2.41, p =. 016). But for presence, the presence of an auditory stimulus had no effect ($\hat{\beta}$ = 0.00, 95% CI [−0.01,0.02], z = 0.47, p =.637). We also found that participants were more confident in their absence than in their presence judgments ($\hat{\beta}$ =−4. 80, 95% CI [−6. 53, −3. 07], t(203.89) =−5.43, p <. 001). They were also more confident in the absence of an auditory stimulus ($\hat{\beta}$ =−2.00, 95% CI [−3.73, −0.28], t(203.40) =−2.28, p =. 024), but the interaction between the presence of an auditory stimulus and the type of judgments was not significant.

Metacognitive efficiency

To further examine metacognitive performance, we analyzed metacognitive efficiency (Mratio) for presence versus absence trials, and across stimulus modalities.

Metacognitive efficiency was higher for presence judgments than absence judgments (ΔM = 0.42, *95% HDI* [0.24, 0.60]). This was also the case at the modality-specific level where participants showed higher auditory metacognitive efficiency when the auditory stimulus was judged present compared to absent(ΔM = 0.34, *95% HDI* [0.12, 0.55]), and higher visual metacognitive efficiency when the visual stimulus was judged present compared to absent (ΔM = 0.33, 95% HDI [0.09, 0.58]).

Moreover, when looking more closely at presence judgments, no credible difference was found between unimodal and bimodal trials (ΔM = -0.09, *95% HDI* [-0.32, 0.14]). There was also no clear difference in metacognitive efficiency between auditory and visual trials (ΔM = -0.06, *95% HDI* [-0.32, 0.19], between auditory and audiovisual trials (ΔM < .001, *95% HDI* [-0.22, 0.23]), or between visual and audiovisual trials (ΔM = -0.06, *95% HDI* [-0.29, 0.16]). Finally, in order to test the supramodality of metaperception, we compute the correlation in metacognitive efficiency between modalities. We found strong evidence for correlation between the audiovisual and auditory trials (*Mcorr* = 0.77, *95% HDI* [0.46, 0.99]), between the audiovisual and visual trials (*Mcorr* = 0.79, *95% HDI* [0.50, 0.99]), and between the auditory and visual trials (*Mcorr* = 0.77, *95% HDI* [0.43, 0.99]).

We also assess auditory metacognitive efficiency as a function of visual information. When the auditory stimulus was judged present, the visual stimulus had no meaningful effect on metacognitive efficiency (ΔM = 0.16, *95% HDI* [-0.17, 0.48]). The visual stimulus also had no impact when the auditory stimulus was judged absent (ΔM = -0.06, *95% HDI* [-0.27, 0.15]).

Finally, we examined visual metacognitive efficiency as a function of the auditory stimulus. There was no clear effect of the auditory stimulus neither when the visual stimulus was judged absent (ΔM = -0.05, 95% HDI [-0.24, 0.15]) nor judged present (ΔM = -0.009, 95% HDI [-0.39, 0.35]).

Data availability

Data, experimental protocol, analysis code, and modelling code used are available at *https://gitlab.com/nfaivre/bimodal_confidence_public*.

Acknowledgements

Funded by the European Union.iews and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them. This work is supported by an ERC grant (Volta, 101125379) awarded to NF. The authors thank the IDEX for funding PP mobility grant.

Additional information

Contributions:

PP: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing - Original, Writing - Review & Editing, Visualisation, Funding acquisition. **MM**: Methodology, Software, Formal analysis, Writing - Review & Editing, Visualisation, Supervision. **CD**: Conceptualization, Methodology, Writing - Review & Editing. **LG**: Conceptualization, Methodology, Validation, Formal analysis, Writing - Review & Editing, Visualisation, Project administration, Supervision. **NF**: Conceptualization, Methodology, Validation, Formal analysis, Resources, Writing - Review & Editing, Visualisation, Project administration, Funding acquisition, Supervision

Funding

Funder	Grant reference number	Author
EC | European Research Council (ERC)	https://doi.org/10.3030/101125379	Nathan Faivre
Idex	UGA	Perrine Porte

References

Aitchison L., Bang D., Bahrami B., Latham P. E. (2015) Doubly Bayesian Analysis of Confidence in Perceptual Decision-Making. *PLOS Computational Biology* **11**:e1004519 https://doi.org/10.1371/journal.pcbi.1004519

Ais J., Zylberberg A., Barttfeld P., Sigman M. (2016) Individual consistency in the accuracy and distribution of confidence judgments. *Cognition* **146**:377-386 https://doi.org/10.1016/j.cognition.2015.10.006 | PubMed

Arbuzova P., Peters C., Röd L., Koß C., Maurer H., Maurer L. K., Müller H., Verrel J., Filevich E. (2021) Measuring metacognition of direct and indirect parameters of voluntary movement. *Journal of Experimental Psychology: General* **150**:2208-2229 https://doi.org/10.1037/xge0000892 | PubMed

Blurton S. P., Greenlee M. W., Gondan M. (2014) Multisensory processing of redundant information in go/no-go and choice responses. *Attention, Perception, & Psychophysics* **76**:1212-1233 https://doi.org/10.3758/s13414-014-0644-0 | PubMed

Callaway F., Griffiths T L., Norman K. A., Zhang Q. (2023) Optimal metacognitive control of memory recall. Psychological Review. https://doi.org/10.1037/rev0000441 | PubMed

Charles L., Chardin C., Haggard P. (2020) Evidence for metacognitive bias in perception of voluntary action. *Cognition* **194**:104041 https://doi.org/10.1016/j.cognition.2019.104041 | PubMed

Deroy O., Spence C., Noppeney U. (2016) Metacognition in Multisensory Perception. *Trends in Cognitive Sciences* **20**:736-747 https://doi.org/10.1016/j.tics.2016.08.006 | PubMed

Dijkstra N., Mazor M., Fleming S. M. (2024) Confidence ratings do not distinguish imagination from reality. *Journal of Vision* **24**:13 https://doi.org/10.1167/jov.24.5.13 | PubMed

Egan J. M., Gomez-Ramirez M., Foxe J. J., O'Connell R. G., Kelly S. P (2025) Distinct audio and visual accumulators co-activate motor preparation for multisensory detection. *Nature Human Behaviour* 1-15 https://doi.org/10.1038/s41562-025-02280-9 | PubMed

Faivre N., Filevich E., Solovey G., Kühn S., Blanke O. (2018) Behavioral, Modeling, and Electrophysiological Evidence for Supramodality in Human Metacoqnition. *Journal of Neuroscience* **38**:263-277 https://doi.org/10.1523/JNEUROSCI.0322-17.2017 | PubMed

Flavell J. H. (1979) Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist* **34**:906-911 https://doi.org/10.1037/0003-066X.34.10.906

Fleming S. M. (2017) HMeta-d: Hierarchical Bayesian estimation of metacognitive efficiency from confidence ratings. *Neuroscience of Consciousness* **2017**:nix007 https://doi.org/10.1093/nc/nix007 | PubMed

Fleming S. M., Dolan R. J., Frith C. D. (2012) Metacognition: Computation, biology and function. *Philosophical Transactions of the Royal Society B: Biological Sciences* **367**:1280-1286 https://doi.org/10.1098/rstb.2012.0021 | PubMed

Fleming S. M., Lau H. C. (2014) Howto measure metacognition. *Frontiers in Human Neuroscience* **8**:443 https://doi.org/10.3389/fnhum.2014.00443 | PubMed

Gondan M., Lange K., Rösler F., Röder B. (2004) The redundant target effect is affected by modality switch costs. *Psychonomic Bulletin & Review* **11**:307-313 https://doi.org/10.3758/BF03196575 | PubMed

Gondan M., Niederhaus B., Rösler F., Röder B. (2005) Multisensory processing in the redundant-target effect: A behavioral and event-related potential study. *Perception & Psychophysics* **67**:713-726 https://doi.org/10.3758/BF03193527 | PubMed

Kellij S., Fahrenfort J., Lau H., Peters M. A. K., Odegaard B. (2020) An investigation of how relative precision of target encoding influences metacognitive performance. *Attention, Perception, & Psychophysics* **83**:512-524 https://doi.org/10.3758/s13414-020-02190-0 | PubMed

Maniscalco B., Lau H. (2012) A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Consciousness and Cognition* **21**:422-430 https://doi.org/10.1016/j.concog.2011.09.021 | PubMed

Mazor M. (2025) Inference About Absence as a Window Into the Mental Self-Model. *Open Mind* **9**:635-651 https://doi.org/10.1162/opmi_a_00206

Mazor M., Fleming S. M. (2020) Distinguishing absence of awareness from awareness of absence. *Philosophy and the Mind Sciences* **1** https://doi.org/10.33735/phimisci.2020.II.69

Mazor M., Friston K. J., Fleming S. M. (2020) Distinct neural contributions to metacognition for detecting, but not discriminating visual stimuli. *eLife* **9**:e53900 https://doi.org/10.7554/eLife.53900 | PubMed

Mazor M., Maimon-Mor R. O., Charles L., Fleming S. M. (2023) Paradoxical evidence weighting in confidence judgments for detection and discrimination. *Attention, Perception, & Psychophysics* **85**:2356-2385 https://doi.org/10.3758/s13414-023-02710-8 | PubMed

Mazor M., Moran R., Press C. (2025) Beliefs about perception shape perceptual inference : An ideal observer model of detection. Psychological Review. https://doi.org/10.1037/rev0000552 | PubMed

Meuwese J. D. I., van Loon A. M., Lamme V. A. F., Fahrenfort J. J. (2014) The subjective experience of object recognition : Comparing metacognition for object detection and object categorization. *Attention, Perception, & Psychophysics* **76**:1057-1068 https://doi.org/10.3758/s13414-014-0643-1 | PubMed

Pereira M., Megevand P., Tan M. X., Chang W., Wang S., Rezai A., Seeck M., Corniola M., Momjian S., Bernasconi F., *et al.* (2021) Evidence accumulation relates to perceptual consciousness and monitoring. *Nature Communications* **12** https://doi.org/10.1038/s41467-021-23540-y | PubMed

Plass J., Brang D. (2021) Multisensory stimuli shift perceptual priors to facilitate rapid behavior. *Scientific Reports* **11** https://doi.org/10.1038/s41598-021-02566-8 | PubMed

Pouget A., Drugowitsch J., Kepecs A. (2016) Confidence and certainty : Distinct probabilistic quantities for different goals. *Nature neuroscience* **19**:366-374 https://doi.org/10.1038/nn.4240 | PubMed

Rach S., Diederich A., Colonius H. (2011) On quantifying multisensory interaction effects in reaction time and detection rate. *Psychological Research* **75**:77-94 https://doi.org/10.1007/s00426-010-0289-0 | PubMed

Rouault M., McWilliams A., Allen M. G., Fleming S. M. (2018) Human Metacognition Across Domains: Insights from Individual Differences and Neuroimaging. *Personality Neuroscience* **1**:e17 https://doi.org/10.1017/pen.2018.16 | PubMed

Schipper M., Mazor M. (2025) Confidence in absence as confidence in counterfactual visibility. *PsyArXiv* https://doi.org/10.31234/osf.io/vuytz_v1

Stockart F., Msheik R., Robin A., Jurkovičová L., Goueytes D., Rouy M., Marecek R., Hoffmann D., Mudrik L., Roman R., *et al.* (2025) Cortical evidence accumulation for visual perception occurs irrespective of reports. *Nature Communications* **16**:8458 https://doi.org/10.1038/s41467-025-63255-y | PubMed

Peer reviews

Reviewer #1 (Public review):

Porte et al. investigate how observers form confidence judgments about the presence vs absence of near-threshold audiovisual stimuli. In two psychophysical detection experiments, human participants judged whether a stimulus (visual, auditory, or audiovisual) was present or absent, reported amodal confidence, and then gave modality-specific detection and confidence ratings using a bidimensional scale. The authors report that audiovisual (AV) stimuli are detected more accurately than unimodal stimuli, but that multisensory stimulation does not improve metacognitive efficiency. Participants are more confident in absence than in presence judgments. They extend a previously proposed model to an audiovisual setting, assuming evidence is available only for presence and that absence is inferred via counterfactual detectability. Detection is modeled with a disjunctive integration rule across modalities, while confidence is explained by a combination of conjunctive (for presence) and disjunctive/negation-of-disjunction (for absence) rules.

There are several points I wish to have clarified, outlined below:

(1) Framing of bimodal vs unimodal detection

On p.3, the introduction states that "Adults typically show higher detection rates and faster reaction times for bimodal than for unimodal stimuli." This is broadly consistent with the literature, but as written, it obscures the fact that these effects depend critically on experimenter-defined stimulus strengths. It is trivial to construct cases where a strong unimodal stimulus is more detectable than a bimodal stimulus made of two very weak unimodal stimuli. If "bimodal" is understood as the co-presentation of two unimodal components matched in detectability, then Bayes-rule-based arguments indeed predict better detection for the bimodal case; how much better is theoretically interesting, but not quantified in this paper. There is an entire literature on the combination of two unimodal stimuli, which is not touched on. For a pertinent reference, see Ernst & Banks 2002. I recommend clarifying that the statement assumes comparable unimodal intensities.

(2) Relationship to signal detection theory and counterfactual perceptibility

In the introduction, the authors write, "If sensory evidence is only available for presence," motivating counterfactual perceptibility as a necessary ingredient to infer absence. However, standard signal detection theory (SDT) already provides a widely accepted framework in which a continuous internal response is present on both signal and noise (absent) trials, with absence corresponding to the noise distribution and decisions implemented by a criterion.

Thus, there is no logical need to invoke counterfactual perceptibility simply to define absence; rather, the Mazor-style framework adds an explicit belief model about detectability and an optimal stopping policy. It would strengthen the paper to more clearly state how the proposed model goes beyond SDT conceptually, acknowledge that SDT can account for presence/absence decisions without counterfactuals, and position the counterfactual account as a hypothesis about how observers actually compute absence/confidence, not as a necessity. One of the central claims of the paper is that detection in the case of absence requires counterfactual reasoning. The authors should demonstrate whether or not an SDT-based generative model can describe these amodal and uni- and bi-modal stimulus decisions. In such an SDT model, an SDT-based generative model in which the noise distribution is shared across conditions, and unimodal vs bimodal differences are captured by changes in the mean or variance of the signal+noise distribution.

(3) Confidence vs performance: is AV confidence special?

The paper's central claims about multisensory confidence and metacognition would be stronger if the authors showed that AV confidence deviates from what is expected given performance alone. From the reported results, AV accuracy is around 80%, with visual and auditory at about 60% and 40%, respectively. Given that confidence typically monotonically scales with accuracy, the first question is whether AV confidence is entirely explained by improved performance, or whether there is an additional multisensory contribution. A simple, informative analysis would be for each subject, plot mean confidence vs per cent correct for AV, V, A, and absent conditions, and to test whether AV confidence lies above the trend predicted by accuracy alone.

(4) Metacognitive measures: logistic regression slopes vs meta-d'/d'

In the "Multisensory effects on metacognitive performance" section, the authors define "metacognitive sensitivity" as the slope of a Bayesian logistic regression predicting accuracy from confidence. There is substantial literature showing that logistic-slope measures of metacognitive sensitivity are criterion-dependent and can be affected by both task and confidence criteria (for one example, see Rausch & Zehetleitner, 2017). In contrast, meta-d'/d' was specifically developed to provide a bias-invariant measure of metacognitive efficiency. Though this, too, is dated (see Boundy-Singer et al., 2023). Given that the authors already estimate HMeta-d-based M-ratios, it is unclear why they rely on logistic regression slopes as their primary "metacognitive sensitivity" metric in Figure 4A. I suggest either replacing the logistic-slope metric with SDT-based measures (meta-d', meta-d'/d') or providing a clear justification for using logistic slopes, along with a discussion of their known limitations.

Additionally, Figure 3 reports M-ratios without showing the corresponding d' or meta-d' for judge-present vs judge-absent conditions. Presenting these would help contextualize the metacognitive efficiency results and clarify whether differences are driven mainly by changes in metacognitive sensitivity, changes in task performance, or both. The d' values per condition could be added to Figure 2A.

(5) Interpretation of confidence in absence vs presence

The authors emphasise that it is surprising subjects are more confident in absence than in presence judgments, both at amodal and modality-specific levels. However, Figure 2B suggests that absent responses are very accurate: absent is reported as present only in about 10% of absent trials, implying a high correct rejection rate. If confidence tracks outcome probability, higher confidence for absence may be at least partly expected. Before attributing this asymmetry primarily to counterfactual reasoning, it would be important to explicitly relate confidence to accuracy for hits, misses, false alarms, and correct rejections and show whether absence confidence remains elevated relative to presence after controlling for accuracy differences across judgment types and conditions. Without this, the interpretation that higher absence confidence is inherently "unexpected" seems overstated.

(6) Model: integration rules, confidence, and evidence strength

The modeling section extends the Mazor et al. ideal observer to two modality-specific sensors, with disjunctive integration for detection and then disjunctive vs conjunctive integration rules for confidence. I have a few comments.

First, the detection rule is disjunctive and is reported as a finding. However, the conclusion that detection relies on a disjunctive rule ("present if A or V") closely mirrors the task instructions-participants are explicitly told to respond "present" if they detect the stimulus in any modality. As such, this seems more like a sanity check than a novel empirical finding.

Relatedly, the conjunctive detection is a weak null. The conjunctive rule ("present only if both A and V") is behaviorally implausible given the task instructions. A more informative baseline would be an SDT-style scalar-evidence model (see comment 2), rather than a conjunctive rule that participants would have to actively violate the instructions to follow.

Second, confidence in the model is defined as the probability of being correct at the time of the detection decision. However, this implies a fixed amount of evidence at decision time unless additional mechanisms are invoked. This issue is well known in diffusion modeling (see Kiani et al. 2014) and deserves explicit discussion; otherwise, it is unclear how the model produces graded confidence from a bound-crossing rule alone.

Third, the authors do not consider a straightforward evidence-strength account of confidence. When both modalities indicate presence, there is, on average, more total sensory evidence than in unimodal trials, making correct decisions more likely and, under most frameworks, confidence higher. Likewise, weak evidence in both modalities can be stronger evidence for absence than moderate in one and weak in the other. Many of the patterns that motivate the presence-conjunctive/absence-disjunctive mix could arise from a model where confidence simply reflects the amount of evidence for the chosen option, without positing distinct logical integration rules for presence vs absence. As the authors note, purely disjunctive or purely conjunctive confidence rules fail to capture the trends in confidence reports in Figure 7, leading them to adopt a combined presence-conjunctive / absence-disjunctive rule. A more parsimonious alternative-that confidence scales with evidence magnitude and cross-modal agreement-should be explicitly considered and, ideally, implemented as a competing model. Finally, if the model is intended as a good account of the data, it would be useful to report whether it also reproduces the metacognitive efficiency patterns (M-ratios) beyond the mean confidence patterns shown in Figures 7-8. At present, the model appears systematically over-confident, which should be acknowledged and quantified.

(7) Confidence asymmetry index (CAI) and modality weighting

The confidence asymmetry index (CAI) is defined as the difference between auditory and visual confidence on AV vs absent trials, and the authors report strong correlations between observed and simulated CAI across participants. They interpret this as evidence that subjects place different weights on auditory vs visual signals. Several questions arise. First, does CAI capture asymmetries beyond what is expected from accuracy differences between modalities and conditions? Second, because the simulated data are generated from model fits to the observed data, a correlation between observed and simulated CAI is expected: the model is built to reproduce the individual patterns it is then compared to. A stronger test would compare CAI from data simulated with modality-specific belief parameters, versus CAI from data simulated with constrained equal belief parameters (same θs). Relatedly, the paper would benefit from a plot showing the distribution of θs for A and V- present stimuli across subjects. These values could also be related to unimodal sensitivity measured in the calibration/training phases. A natural prediction is that higher unimodal sensitivity should correspond to higher belief parameters for presence.

https://doi.org/10.7554/eLife.110765.1.sa3

Reviewer #2 (Public review):

Summary:

In this study, across two experiments, the authors wrestle with the question: What is the profile of confidence judgments in presence/absence decisions for audiovisual stimuli? After thresholding observers to 50% target detection rates in each modality, the authors conducted one experiment that included 75% target presence (spread equally across bimodal, auditory, and visual targets) and one experiment with 50% overall target presence. Results showed that, overall, detection performance was higher for audiovisual stimuli compared to unimodal ones, and that a recent model for stimulus detection could be extended to this multisensory scenario. By incorporating a disjunctive rule for absence judgments and a conjunctive rule for presence judgments, the model was able to qualitatively reproduce some of the trends observed in the human data regarding confidence.

Strengths:

(1) The paper makes novel contributions to the study of multisensory confidence judgments for yes/no target detection.

(2) The paper further extends the use of a leading model of stimulus detection (from Mazor et al., 2025).

(3) Pre-registration of the study was implemented, and the code is publicly available (although the GitLab link requires registration to access the materials).

(4) One of the empirical results (higher confidence for absence compared to presence judgments) is especially interesting, contributing another empirical finding to a very mixed literature on this topic (as the authors note).

Weaknesses:

(1) Page 5 - I have concerns about the use of the equal-variance model from Signal Detection Theory to analyze the data. For example, the authors should read the recent paper by Miyoshi, Rahnev, and Lau in iScience, found at this link: *https://www.cell.com/iscience/fulltext/S2589-0042(26)00373-1*. In this paper, the authors note how the equal variance model should be used with caution in yes/no detection tasks, since the variances of the "stimulus present" and "stimulus absent" distributions are often different from one another. In a revision, I highly recommend that the authors explicitly discuss this paper and review whether the assumptions for the equal-variance model have been met (e.g., since they have confidence data, one way to do this would be to evaluate if the slope of the line in zROC space differs from 1). The authors may also want to incorporate methods from this iScience paper into the current manuscript, or potentially move to using an unequal variance SDT model and compute d'a and c'a.

(2) Related to the computation/measurement of the response criterion, the authors note on page 18 in the Methods that for Experiment 1, signals are actually present on 75% of trials, since a bimodal stimulus is present on 25% of trials, the visual circle only occurs on 25% of trials, the sinusoidal tone occurs on 25% of trials, and then only noise is present on 25% of trials. Did the authors have any a priori hypotheses about the response criteria that participants would exhibit in Experiment 1, considering the unbalanced target presentation rate in this task? Also, in Experiment 2, what did it mean to equate target present and target absent trials? Is it that they broke 50% target present trials down into 16.67% bimodal targets,

16.67% visual targets, and 16.67% auditory targets? A few more details would be good to explicitly note for those trying to replicate the task.

(3) It is important to plot the individual data for Figure 2. If the authors didn't match detection performance for the visual and auditory modalities, it would be good to see the individual data to know why. Is it that the thresholding procedure didn't work for some of the participants in the visual modality, and that's why the "yes" response rate is (on average) ~60% or higher across the two experiments? Similarly, in the auditory domain, do the authors have participants that are at floor? Or is it simply that the staircases failed to successfully target 50% detection on average?

(4) The authors mentioned that data were collected on the Prolific platform. What checks did they conduct to ensure that this data wasn't produced by bots? There are recent high-profile publications in PNAS and Behavioral Research Methods that indicate how online data collection is problematic (e.g., *https://www.pnas.org/doi/10.1073/pnas.2535585123* and *https://link.springer.com/article/10.3758/s13428-025-02852-7*). What analyses or quality checks are there to ensure that humans were the ones completing the task?

(5) Page 7 - Since confidence was collected on a continuous scale, the authors should say a bit more about how they were able to compute measures of metacognitive efficiency. My understanding is that to compute meta-d', the data has to be binned. How was the binning implemented? With whatever bin size the authors chose, would it make any difference to the results if they changed the number of the bins in the analysis?

(6) Page 8 - Is there a prior precedent for using slope of the Bayesian logistic regression predicting accuracy from confidence as a measure of metacognitive sensitivity? If so, can the authors cite those papers as a reference? If not, can they place this analysis within the context of other measures of metacognitive sensitivity that exist? (meta-d', AUROC (Type 2), etc.)

(7) Page 8 - Another one of the results on page 8 is worth reflecting further upon: the authors note how in Experiment 1, no credible difference was found between unimodal and bimodal trials (DeltaM = -0.25 [-0.59, 0.10]), but in Experiment 2, "we observed higher metacognitive efficiency in unimodal compared to bimodal trials (DeltaM = -0.28 [-0.54, -0.02]. Those DeltaM values are nearly identical, so without a power analysis motivating the number of participants the authors collected, how certain are they that the results from these two experiments are really that distinct? It reminds me a bit of the Andrew Gelman blog post, "The difference between significance and non-significance is not significant".

(8) Is there any way to look at whether the presence of multisensory hallucinations (or perhaps that word is too strong, and we should simply consider them miscategorizations) increased as the task progressed? That is, the authors have repeated presentations of audiovisual stimuli for at least some percentage of the trials. Since the percentages for auditory stimuli being correctly categorized as auditory are at 85% in Experiment 1 and 79% in Experiment 2, were the trials where they miscategorized these stimuli equally spread throughout the task? Or did they come later in the experiment, after being repeatedly exposed to multisensory trials?

(9) Would the authors obtain the same results if they got rid of the amodal confidence judgment in their task, and simply had participants report the bimodal confidence following the presence/absence judgment? Part of the reason for asking this is that, according to page 11, the model is only fitted to amodal detection accuracy and response time data. This surprised me. I would have expected that the bimodal confidence would provide more useful information for the model fit. The authors should further explain this rationale in the paper. It seems odd to me to have the multisensory confidence ratings and not have them play a central role in the modeling work.

(10) In Figure 6, it appears the model is a bit off in its estimate of auditory responses (panel B, E) in the AV condition. Do the authors have any intuitions about why this might be happening?

(11) The authors talk about how the model is reproducing effects in the human data, but there's no systematic comparison, quantitatively, of how the two things relate. The authors should include some quantitative measure that reflects this.

(12) Related to this, I am not sure I agree with the characterization in Figure 7 that "when confidence followed a disjunctive rule, the model failed to capture important aspects of the data. On the other hand, when confidence followed a conjunctive rule, it reproduced confidence in presence judgments but failed to capture variability in confidence ratings for absence judgments." What, quantitatively, is the basis of this claim? This applies to Figure 8, too. I am not clear how, specifically, and quantitatively, the authors are justifying their claims about model fits. I don't think the confidence asymmetry index in Figure 8 is enough to quantify the quality of the model fitting procedure.

(13) Is there any chance the higher metacognitive efficiency for auditory trials is simply driven by differences in the d' values across the modalities? It might be good to probe this effect further.

(14) Lastly, I think it would be interesting to look at how instructions about modality-specific attention could modulate these findings, in terms of how unimodal (unimodal visual, unimodal auditory) or bimodal attention might modulate these effects. This is an idea for future work.

https://doi.org/10.7554/eLife.110765.1.sa2

Reviewer #3 (Public review):

Summary:

This study used a pre-registered novel behavioural paradigm and computational modelling to investigate multi-sensory influences on detection and confidence. Participants performed amodal detection of auditory and visual stimuli (indicating that a stimulus was there when either an auditory stimulus or a visual stimulus or both were present), followed by amodal and unimodal confidence ratings. Detection was higher when both stimuli were present, and the presence of one modality increased the confidence in the presence of the other modality. In contrast to previous detection studies, confidence was higher for absent than for present judgements, but metacognitive efficiency was higher for present judgements. Metacognitive sensitivity was higher for bimodal stimuli, but this was not the case for metacognitive efficiency, suggesting that the sensitivity might be driven by first-order performance. The computational model showed that both detection and confidence in absence followed a disjunctive evidence integration rule, while confidence in presence followed a conjunctive integration rule.

Strengths:

The paper has several major strengths. Firstly, it addresses a novel research question using an innovative and well-controlled paradigm. Furthermore, the paradigm and analyses were pre-registered, and all effects that were interpreted were replicated in two independent samples. Finally, the paper uses an advanced computational model to capture counterintuitive patterns in the data.

Weaknesses:

The major weakness of the paper is the narrative structure. It is not always clear how the different analyses relate to the main research question. Many different effects are reported in terms of detection accuracy, bias, confidence and metacognition, as well as cross-modal and unimodal versus bimodal effects. It would help readability if the paper were streamlined in terms of the research question that is being answered, which I believe is specifically about multimodal absence judgements. Relatedly, for a reader not intimately familiar with the metacognition literature, the difference between MRatio, metacognitive sensitivity and metacognitive efficiency is not obvious. It would be good to clarify this more in the manuscript.

In general, the conclusions drawn by the authors seem to be supported by the results. However, I was missing quantitative model comparisons between the conjunctive and the disjunctive models and an explanation of why the models systematically overestimated the confidence ratings. Furthermore, the 'perceptual multisensory interference' section reports on very interesting effects, but these are not supported by statistical tests in the main text. It would help to assess the strength of the claims if the statistical evidence in favour of these claims were presented together in the main text.

One other concern is that in real-world multi-sensory perception, such as the mosquito example in the introduction, the auditory and visual signals have a strong natural association, which means that if you hear the auditory signal, you expect that you will see the visual signal soon and vice versa. As far as I understood, this association was not present in the current paradigm, which might influence the type of effects that one would expect to see.

https://doi.org/10.7554/eLife.110765.1.sa1

Author response:

> ***Reviewer 1:***
>
> *Porte et al. investigate how observers form confidence judgments about the presence vs absence of near-threshold audiovisual stimuli. In two psychophysical detection experiments, human participants judged whether a stimulus (visual, auditory, or audiovisual) was present or absent, reported amodal confidence, and then gave modality-specific detection and confidence ratings using a bidimensional scale. The authors report that audiovisual (AV) stimuli are detected more accurately than unimodal stimuli, but that multisensory stimulation does not improve metacognitive efficiency. Participants are more confident in absence than in presence judgments. They extend a previously proposed model to an audiovisual setting, assuming evidence is available only for presence and that absence is inferred via counterfactual detectability. Detection is modeled with a disjunctive integration rule across modalities, while confidence is explained by a combination of conjunctive (for presence) and disjunctive/negation-of-disjunction (for absence) rules.*

We thank the reviewer for thoroughly evaluating our work.

> *There are several points I wish to have clarified, outlined below:*
>
> *(1) Framing of bimodal vs unimodal detection*
>
> *On p.3, the introduction states that "Adults typically show higher detection rates and faster reaction times for bimodal than for unimodal stimuli." This is broadly consistent with the literature, but as written, it obscures the fact that these effects depend critically on experimenter-defined stimulus strengths. It is trivial to construct cases where a strong unimodal stimulus is more detectable than a bimodal stimulus made of two very weak unimodal stimuli. If "bimodal" is understood as the co-presentation of two unimodal*

> *components matched in detectability, then Bayes-rule-based arguments indeed predict better detection for the bimodal case; how much better is theoretically interesting, but not quantified in this paper. There is an entire literature on the combination of two unimodal stimuli, which is not touched on. For a pertinent reference, see Ernst & Banks 2002. I recommend clarifying that the statement assumes comparable unimodal intensities.*

We will clarify that when discussing bimodal stimuli, we mean the co-presentation of two unimodal stimuli of similar intensity. We will add references to the literature during discrimination tasks that have shown that multisensory cue-combination followed Bayes rule integration (e.g., Ernst & Banks, 2002; Battaglia et al., 2003; Alais & Burr, 2004) and clarify in which ways our work differs from this rich body of work and provides novel contributions.

> *(2) Relationship to signal detection theory and counterfactual perceptibility*
>
> *In the introduction, the authors write, "If sensory evidence is only available for presence," motivating counterfactual perceptibility as a necessary ingredient to infer absence. However, standard signal detection theory (SDT) already provides a widely accepted framework in which a continuous internal response is present on both signal and noise (absent) trials, with absence corresponding to the noise distribution and decisions implemented by a criterion. Thus, there is no logical need to invoke counterfactual perceptibility simply to define absence; rather, the Mazor-style framework adds an explicit belief model about detectability and an optimal stopping policy. It would strengthen the paper to more clearly state how the proposed model goes beyond SDT conceptually, acknowledge that SDT can account for presence/absence decisions without counterfactuals, and position the counterfactual account as a hypothesis about how observers actually compute absence/confidence, not as a necessity.*
>
> *One of the central claims of the paper is that detection in the case of absence requires counterfactual reasoning. The authors should demonstrate whether or not an SDT-based generative model can describe these amodal and uni- and bi-modal stimulus decisions. In such an SDT model, an SDT-based generative model in which the noise distribution is shared across conditions, and unimodal vs bimodal differences are captured by changes in the mean or variance of the signal+noise distribution.*

We will clarify that our framework explains how absence judgments (and related confidence) are formed, and what it adds to SDT models, including the reproduction of reaction times and a normative explanation of criterion placement (results about RTs are available in the supplementary materials).We will also run additional model comparisons assessing how an SDT-based generative model performs compared to our Bayesian model based on counterfactual perceivability.

> *(3) Confidence vs performance: is AV confidence special?*
>
> *The paper's central claims about multisensory confidence and metacognition would be stronger if the authors showed that AV confidence deviates from what is expected given performance alone. From the reported results, AV accuracy is around 80%, with visual and auditory at about 60% and 40%, respectively. Given that confidence typically monotonically scales with accuracy, the first question is whether AV confidence is entirely explained by improved performance, or whether there is an additional multisensory contribution. A simple, informative analysis would be for each subject, plot mean confidence vs per cent correct for AV, V, A, and absent conditions, and to test whether AV confidence lies above the trend predicted by accuracy alone.*

This is an excellent suggestion, and we will conduct the proposed analysis.

> *(4) Metacognitive measures: logistic regression slopes vs meta-d'/d'*

> *In the "Multisensory effects on metacognitive performance" section, the authors define "metacognitive sensitivity" as the slope of a Bayesian logistic regression predicting accuracy from confidence. There is substantial literature showing that logistic-slope measures of metacognitive sensitivity are criterion-dependent and can be affected by both task and confidence criteria (for one example, see Rausch & Zehetleitner, 2017). In contrast, meta-d'/d' was specifically developed to provide a bias-invariant measure of metacognitive efficiency. Though this, too, is dated (see Boundy-Singer et al., 2023). Given that the authors already estimate HMeta-d-based M-ratios, it is unclear why they rely on logistic regression slopes as their primary "metacognitive sensitivity" metric in Figure 4A. I suggest either replacing the logistic-slope metric with SDT-based measures (meta-d', meta-d'/d') or providing a clear justification for using logistic slopes, along with a discussion of their known limitations.*
>
> *Additionally, Figure 3 reports M-ratios without showing the corresponding d' or meta-d' for judge-present vs judge-absent conditions. Presenting these would help contextualize the metacognitive efficiency results and clarify whether differences are driven mainly by changes in metacognitive sensitivity, changes in task performance, or both. The d' values per condition could be added to Figure 2A.*

All typical measures of metacognitive sensitivity are influenced by metacognitive bias and task performance to some extent, and none of them is a pure measure of type-2 sensitivity (e.g., see Rahnev, 2025). Here, we chose logistic regression because it enables modeling interactions with other predictors in a factorial design with a limited number of trials.

We will clarify the limitations of metacognitive sensitivity measures and better explain why we then used Mratio to estimate metacognitive performance while controlling for underlying task performance.

Thank you for this suggestion. We will add the d’ values per condition to Figure 2A.

> *(5) Interpretation of confidence in absence vs presence*
>
> *The authors emphasise that it is surprising subjects are more confident in absence than in presence judgments, both at amodal and modality-specific levels. However, Figure 2B suggests that absent responses are very accurate: absent is reported as present only in about 10% of absent trials, implying a high correct rejection rate. If confidence tracks outcome probability, higher confidence for absence may be at least partly expected. Before attributing this asymmetry primarily to counterfactual reasoning, it would be important to explicitly relate confidence to accuracy for hits, misses, false alarms, and correct rejections and show whether absence confidence remains elevated relative to presence after controlling for accuracy differences across judgment types and conditions. Without this, the interpretation that higher absence confidence is inherently "unexpected" seems overstated.*

This higher confidence for absence judgments than for presence judgments was observed while controlling for response accuracy. We will clarify this in the main text.

> *(6) Model: integration rules, confidence, and evidence strength*
>
> *The modeling section extends the Mazor et al. ideal observer to two modality-specific sensors, with disjunctive integration for detection and then disjunctive vs conjunctive integration rules for confidence. I have a few comments.*
>
> *First, the detection rule is disjunctive and is reported as a finding. However, the conclusion that detection relies on a disjunctive rule ("present if A or V") closely mirrors the task instructions-participants are explicitly told to respond "present" if they detect the*

stimulus in any modality. As such, this seems more like a sanity check than a novel empirical finding. Relatedly, the conjunctive detection is a weak null. The conjunctive rule ("present only if both A and V") is behaviorally implausible given the task instructions. A more informative baseline would be an SDT-style scalar-evidence model (see comment 2), rather than a conjunctive rule that participants would have to actively violate the instructions to follow.

Second, confidence in the model is defined as the probability of being correct at the time of the detection decision. However, this implies a fixed amount of evidence at decision time unless additional mechanisms are invoked. This issue is well known in diffusion modeling (see Kiani et al. 2014) and deserves explicit discussion; otherwise, it is unclear how the model produces graded confidence from a bound-crossing rule alone.

Third, the authors do not consider a straightforward evidence-strength account of confidence. When both modalities indicate presence, there is, on average, more total sensory evidence than in unimodal trials, making correct decisions more likely and, under most frameworks, confidence higher. Likewise, weak evidence in both modalities can be stronger evidence for absence than moderate in one and weak in the other. Many of the patterns that motivate the presence-conjunctive/absence-disjunctive mix could arise from a model where confidence simply reflects the amount of evidence for the chosen option, without positing distinct logical integration rules for presence vs absence. As the authors note, purely disjunctive or purely conjunctive confidence rules fail to capture the trends in confidence reports in Figure 7, leading them to adopt a combined presence-conjunctive/absence-disjunctive rule. A more parsimonious alternative-that confidence scales with evidence magnitude and cross-modal agreement-should be explicitly considered and, ideally, implemented as a competing model. Finally, if the model is intended as a good account of the data, it would be useful to report whether it also reproduces the metacognitive efficiency patterns (M-ratios) beyond the mean confidence patterns shown in Figures 7-8. At present, the model appears systematically over-confident, which should be acknowledged and quantified.

Indeed, the disjunctive rule was expected, given our design; we will clarify this. As mentioned above, we will directly compare the results of our current model with those of a more traditional SDT-based generative model, as suggested by the reviewer.

Contrary to a classical drift diffusion model, the model does not assume a fixed decision boundary, but derives an optimal stopping policy per time point and belief state. As a result, and depending on beliefs about perceptual evidence and the temporal discounting factor, optimal decision boundaries can be asymmetric and may collapse asymmetrically toward 0. Furthermore, given the asymmetry in the information value between sensor activations and inactivations, and differences in the information value of sensor activations of the two modalities, boundary crossing can lead to belief states that are far or close to the decision boundary, depending on the nature of the evidence. Together, even without an explicit modeling of post-decisional evidence, the model can account for variability in the total accumulated evidence at decision time.

From our understanding, the proposed alternative is equivalent to our current model, in which confidence scales with evidence magnitude.

The model was not fitted to confidence data, which could explain its overall overconfidence. To further test our model, we will assess its ability to reproduce patterns of metacognitive efficiency (M-ratios).

(7) Confidence asymmetry index (CAI) and modality weighting

The confidence asymmetry index (CAI) is defined as the difference between auditory and *visual confidence on AV vs absent trials, and the authors report strong correlations between observed and simulated CAI across participants. They interpret this as evidence that subjects place different weights on auditory vs visual signals. Several questions arise. First, does CAI capture asymmetries beyond what is expected from accuracy differences between modalities and conditions? Second, because the simulated data are generated from model fits to the observed data, a correlation between observed and simulated CAI is expected: the model is built to reproduce the individual patterns it is then compared to. A stronger test would compare CAI from data simulated with modality-specific belief parameters, versus CAI from data simulated with constrained equal belief parameters (same θs). Relatedly, the paper would benefit from a plot showing the distribution of θs for A and V- present stimuli across subjects. These values could also be related to unimodal sensitivity measured in the calibration/training phases. A natural prediction is that higher unimodal sensitivity should correspond to higher belief parameters for presence.*

The model was not fitted to either the modality-specific responses or the confidence ratings, so the correlation between observed and simulated CAI was not expected and provides a good test of our model's ability to reproduce the observed patterns. We will test whether the same correlations hold when using the difference in accuracy instead of the confidence.

We found that the best model is the one with the same belief across the visual and auditory sensors. Given this, we cannot investigate how modality-specific belief parameters are linked to unimodal sensitivity for each participant.

Reviewer 2:

Summary:

In this study, across two experiments, the authors wrestle with the question: What is the profile of confidence judgments in presence/absence decisions for audiovisual stimuli? After thresholding observers to 50% target detection rates in each modality, the authors conducted one experiment that included 75% target presence (spread equally across bimodal, auditory, and visual targets) and one experiment with 50% overall target presence. Results showed that, overall, detection performance was higher for audiovisual stimuli compared to unimodal ones, and that a recent model for stimulus detection could be extended to this multisensory scenario. By incorporating a disjunctive rule for absence judgments and a conjunctive rule for presence judgments, the model was able to qualitatively reproduce some of the trends observed in the human data regarding confidence.

Strengths:

(1) The paper makes novel contributions to the study of multisensory confidence judgments for yes/no target detection.

(2) The paper further extends the use of a leading model of stimulus detection (from Mazor et al., 2025).

(3) Pre-registration of the study was implemented, and the code is publicly available (although the GitLab link requires registration to access the materials).

(4) One of the empirical results (higher confidence for absence compared to presence judgments) is especially interesting, contributing another empirical finding to a very mixed literature on this topic (as the authors note).

We thank the reviewer for the positive evaluation of our work.

> *Weaknesses:*
>
> *(1) Page 5 - I have concerns about the use of the equal-variance model from Signal Detection Theory to analyze the data. For example, the authors should read the recent paper by Miyoshi, Rahnev, and Lau in iScience, found at this link: https://www.cell.com/iscience/fulltext/S2589-0042(26)00373-1 . In this paper, the authors note how the equal variance model should be used with caution in yes/no detection tasks, since the variances of the "stimulus present" and "stimulus absent" distributions are often different from one another. In a revision, I highly recommend that the authors explicitly discuss this paper and review whether the assumptions for the equal-variance model have been met (e.g., since they have confidence data, one way to do this would be to evaluate if the slope of the line in zROC space differs from 1). The authors may also want to incorporate methods from this iScience paper into the current manuscript, or potentially move to using an unequal variance SDT model and compute d'a and c'a.*

This is an excellent suggestion. We will run this analysis and refit the d' and criterion response using unequal-variance models to see whether we observe the same results.

> *(2) Related to the computation/measurement of the response criterion, the authors note on page 18 in the Methods that for Experiment 1, signals are actually present on 75% of trials, since a bimodal stimulus is present on 25% of trials, the visual circle only occurs on 25% of trials, the sinusoidal tone occurs on 25% of trials, and then only noise is present on 25% of trials. Did the authors have any a priori hypotheses about the response criteria that participants would exhibit in Experiment 1, considering the unbalanced target presentation rate in this task? Also, in Experiment 2, what did it mean to equate target present and target absent trials? Is it that they broke 50% target present trials down into 16.67% bimodal targets, 16.67% visual targets, and 16.67% auditory targets? A few more details would be good to explicitly note for those trying to replicate the task*

We will clarify this point in the manuscript. In Experiment 2, the stimulus was absent on 50% of the trials. As a result, the 50% of stimulus present trials were split into the three possible conditions, resulting in a sixth of the trials being auditory, a sixth visual, and a sixth audiovisual; we will make these proportions clearer in the text.

We did not have any a priori hypotheses about the response criteria for Experiment 1. The reviewer is right, the proportion of absent versus present trials can indeed have an impact on response bias. In fact, one of the goals of Experiment 2 was to test whether the low frequency of absent trials compared to present ones could explain both response bias and higher confidence in absence observed in Experiment 1, which we found was not the case, as we did not observe a difference between the two experiments. We will clarify this in our revision.

> *(3) It is important to plot the individual data for Figure 2. If the authors didn't match detection performance for the visual and auditory modalities, it would be good to see the individual data to know why. Is it that the thresholding procedure didn't work for some of the participants in the visual modality, and that's why the "yes" response rate is (on average) ~60% or higher across the two experiments? Similarly, in the auditory domain, do the authors have participants that are at floor? Or is it simply that the staircases failed to successfully target 50% detection on average?*

We will add individual data to Figure 2.

Indeed, staircases failed to achieve 50% detection on average; participants for whom psychometric curves did not converge were excluded, as were those at floor level in one of the two modalities.

> *(4) The authors mentioned that data were collected on the Prolific platform. What checks did they conduct to ensure that this data wasn't produced by bots? There are recent high-profile publications in PNAS and Behavioral Research Methods that indicate how online data collection is problematic (e.g., https://www.pnas.org/doi/10.1073/pnas.2535585123and https://link.springer.com/article/10.3758/s13428-025-02852-7). What analyses or quality checks are there to ensure that humans were the ones completing the task?*

Data were collected on the Prolific platform, which has been shown to yield high-quality data (Kay, 2025). However, we agree that this is a potential concern and will add a note of caution in the revised manuscript, even if the risk that the data do not come from humans but from bots is low (Huskey et al., 2026; Chetverikov, 2026).

> *(5) Page 7 - Since confidence was collected on a continuous scale, the authors should say a bit more about how they were able to compute measures of metacognitive efficiency. My understanding is that to compute meta-d', the data has to be binned. How was the binning implemented? With whatever bin size the authors chose, would it make any difference to the results if they changed the number of the bins in the analysis?*

We will clarify this aspect of the analysis. Data were binned into four quartiles based on the overall distribution of confidence values across participants, based on the binning used in the example in Fleming (2017). We will examine whether changing the number of bins changes the results (Dayan, 2023).

> *(6) Page 8 - Is there a prior precedent for using slope of the Bayesian logistic regression predicting accuracy from confidence as a measure of metacognitive sensitivity? If so, can the authors cite those papers as a reference? If not, can they place this analysis within the context of other measures of metacognitive sensitivity that exist? (meta-d', AUROC (Type 2), etc.)*

Yes, logistic regression has been used to quantify metacognitive sensitivity before. We will add the relevant papers as references (e.g., Sandberg et al., 2010; Norman et al., 2011; Siedlecka et al., 2016; Wierzchoń et al., 2012; Faivre et al., 2018; Pereira et al., 2023)

> *(7) Page 8 - Another one of the results on page 8 is worth reflecting further upon: the authors note how in Experiment 1, no credible difference was found between unimodal and bimodal trials (DeltaM = -0.25 [-0.59, 0.10]), but in Experiment 2, "we observed higher metacognitive efficiency in unimodal compared to bimodal trials (DeltaM = -0.28 [-0.54, -0.02]. Those DeltaM values are nearly identical, so without a power analysis motivating the number of participants the authors collected, how certain are they that the results from these two experiments are really that distinct? It reminds me a bit of the Andrew Gelman blog post, "The difference between significance and non-significance is not significant".*

The number of participants was determined using a Bayesian optional stopping rule, as preregistered. The reviewer is right that the delta values are very similar in the two experiments. Given that a difference was found in only one experiment, we decided not to draw conclusions from it.

> *(8) Is there any way to look at whether the presence of multisensory hallucinations (or perhaps that word is too strong, and we should simply consider them miscategorizations) increased as the task progressed? That is, the authors have repeated presentations of audiovisual stimuli for at least some percentage of the trials. Since the percentages for auditory stimuli being correctly categorized as auditory are at 85% in Experiment 1 and 79% in Experiment 2, were the trials where they miscategorized these*

> *stimuli equally spread throughout the task? Or did they come later in the experiment, after being repeatedly exposed to multisensory trials?*

We will examine how the proportion of miscategorisation changed throughout the task.

> *(9) Would the authors obtain the same results if they got rid of the amodal confidence judgment in their task, and simply had participants report the bimodal confidence following the presence/absence judgment? Part of the reason for asking this is that, according to page 11, the model is only fitted to amodal detection accuracy and response time data. This surprised me. I would have expected that the bimodal confidence would provide more useful information for the model fit. The authors should further explain this rationale in the paper. It seems odd to me to have the multisensory confidence ratings and not have them play a central role in the modeling work.*

Our main goal was to investigate how participants form integrated, supramodal confidence judgments on the basis of multisensory sources of information. Therefore, the amodal confidence judgments are required here.

Moreover, the model was fitted to response times that corresponded to the amodal judgment. Because we had no meaningful response times for the modality-specific judgment, we could not use them to fit the model.

> *(10) In Figure 6, it appears the model is a bit off in its estimate of auditory responses (panel B, E) in the AV condition. Do the authors have any intuitions about why this might be happening?*

Indeed, the model does not capture the full behavioral effects reflecting multisensory interference in the modality-specific responses. We suppose that the model does not reproduce these interferences, as it is only fitted to amodal detection accuracy, and as the two sensors are completely independent from one another. We will clarify this aspect in the text.

> *(11) The authors talk about how the model is reproducing effects in the human data, but there's no systematic comparison, quantitatively, of how the two things relate. The authors should include some quantitative measure that reflects this*

In addition to the d' and criterion comparison between the observed and simulated data, we will compare modality-specific d' and the correlations between observed and simulated confidence.

> *(12) Related to this, I am not sure I agree with the characterization in Figure 7 that "when confidence followed a disjunctive rule, the model failed to capture important aspects of the data. On the other hand, when confidence followed a conjunctive rule, it reproduced confidence in presence judgments but failed to capture variability in confidence ratings for absence judgments." What, quantitatively, is the basis of this claim? This applies to Figure 8, too. I am not clear how, specifically, and quantitatively, the authors are justifying their claims about model fits. I don't think the confidence asymmetry index in Figure 8 is enough to quantify the quality of the model fitting procedure.*

To further support this claim, we will add a quantitative comparison of the different confidence fits.

> *(13) Is there any chance the higher metacognitive efficiency for auditory trials is simply driven by differences in the d' values across the modalities? It might be good to probe this effect further.*

Thank you for this remark. Indeed, the difference in metacognitive efficiency may be driven by differences in the d' values, and so a lower d' for auditory stimuli can lead to higher metacognitive efficiency for a similar metacognitive sensitivity.

> ***Reviewer 3:***
>
> *This study used a pre-registered novel behavioural paradigm and computational modelling to investigate multi-sensory influences on detection and confidence. Participants performed amodal detection of auditory and visual stimuli (indicating that a stimulus was there when either an auditory stimulus or a visual stimulus or both were present), followed by amodal and unimodal confidence ratings. Detection was higher when both stimuli were present, and the presence of one modality increased the confidence in the presence of the other modality. In contrast to previous detection studies, confidence was higher for absent than for present judgements, but metacognitive efficiency was higher for present judgements. Metacognitive sensitivity was higher for bimodal stimuli, but this was not the case for metacognitive efficiency, suggesting that the sensitivity might be driven by first-order performance. The computational model showed that both detection and confidence in absence followed a disjunctive evidence integration rule, while confidence in presence followed a conjunctive integration rule.*

We thank the reviewer for engaging with our work.

> *Strengths:*
>
> *The paper has several major strengths. Firstly, it addresses a novel research question using an innovative and well-controlled paradigm. Furthermore, the paradigm and analyses were pre-registered, and all effects that were interpreted were replicated in two independent samples. Finally, the paper uses an advanced computational model to capture counterintuitive patterns in the data.*
>
> *Weaknesses:*
>
> *The major weakness of the paper is the narrative structure. It is not always clear how the different analyses relate to the main research question. Many different effects are reported in terms of detection accuracy, bias, confidence and metacognition, as well as cross-modal and unimodal versus bimodal effects. It would help readability if the paper were streamlined in terms of the research question that is being answered, which I believe is specifically about multimodal absence judgements. Relatedly, for a reader not intimately familiar with the metacognition literature, the difference between MRatio, metacognitive sensitivity and metacognitive efficiency is not obvious. It would be good to clarify this more in the manuscript.*

We will improve the narrative structure so that each result clearly relates to the research question.

We will also add a clearer definition of the various metacognition metrics to improve readability.

> *In general, the conclusions drawn by the authors seem to be supported by the results. However, I was missing quantitative model comparisons between the conjunctive and the disjunctive models and an explanation of why the models systematically overestimated the confidence ratings. Furthermore, the 'perceptual multisensory interference' section reports on very interesting effects, but these are not supported by statistical tests in the main text. It would help to assess the strength of the claims if the statistical evidence in favour of these claims were presented together in the main text.*

The model was not fitted to confidence data, which could explain its overall overconfidence. As stated in previous responses, we will perform additional analyses to evaluate the model's ability to reproduce confidence ratings. As some of the results were not replicated across experiments, we decided to put all statistical results related to multisensory interference in the supplementary materials and to focus only on consistent results across experiments.

> *One other concern is that in real-world multi-sensory perception, such as the mosquito example in the introduction, the auditory and visual signals have a strong natural association, which means that if you hear the auditory signal, you expect that you will see the visual signal soon and vice versa. As far as I understood, this association was not present in the current paradigm, which might influence the type of effects that one would expect to see.*

The relation here is indeed artificial; we try to reinforce it as much as possible in the instructions of the task by indicating to the participants that they have to "detect a mosquito" that could be present auditory, visually, or both. But we acknowledge that the association between the visual and auditory stimuli is artificial, which may indeed influence our results.

References

Alais, D., & Burr, D. (2004). The Ventriloquist Effect Results from Near-Optimal Bimodal Integration. Current Biology, 14(3), 257- 262. *https://doi.org/10.1016/j.cub.2004.01.029*

Battaglia, P. W., Jacobs, R. A., & Aslin, R. N. (2003). Bayesian integration of visual and auditory signals for spatial localization. JOSA A, 20(7), 1391- 1397. *https://doi.org/10.1364/JOSAA.20.001391*

Chetverikov, A. (2026). Online behavioral studies are safe for now : Unusual RTs do not imply bots (A reply to Van der Stigchel et al., 2026) (Gjw5u_v1). PsyArXiv. *https://osf.io/preprints/psyarxiv/gjw5u_v1/*

Dayan P. (2023). Metacognitive Information Theory. Open mind : discoveries in cognitive science, 7, 392–411. *https://doi.org/10.1162/opmi_a_00091*

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. Nature, 415(6870), Article 6870. *https://doi.org/10.1038/415429a*

Faivre, N., Filevich, E., Solovey, G., Kühn, S., & Blanke, O. (2018). Behavioral, Modeling, and Electrophysiological Evidence for Supramodality in Human Metacognition. Journal of Neuroscience, 38(2), 263- 277. *https://doi.org/10.1523/JNEUROSCI.0322-17.2017*

Fleming, S. M. (2017). HMeta-d : Hierarchical Bayesian estimation of metacognitive efficiency from confidence ratings. Neuroscience of Consciousness, 2017(1),

Huskey, R., Zhao, Z., Parry, D. A., & Fisher, J. T. (2026). An AI agent can complete the Attention Network Test with human-like behavioral signatures : Implications for the bot-or-not debate (T2jru_v1). PsyArXiv. *https://osf.io/preprints/psyarxiv/t2jru_v1/*

Kay, C.S. Why you shouldn't trust data collected on MTurk. Behav Res 57, 340 (2025). *https://doi.org/10.3758/s13428-025-02852-7nix007* . *https://doi.org/10.1093/nc/nix007*

Norman, E., Price, M. C., & Jones, E. (2011). Measuring strategic control in artificial grammar learning. Consciousness and Cognition, 20(4), 1920-1929. *https://doi.org/10.1016/j.concog.2011.07.008*

Pereira, M., Skiba, R., Cojan, Y., Vuilleumier, P., & Bègue, I. (2023). Preserved Metacognition for Undetected Visuomotor Deviations. Journal of Neuroscience, 43(35), 6176- 6184.

https://doi.org/10.1523/JNEUROSCI.0133-23.2023

Rahnev, D. (2025). A comprehensive assessment of current methods for measuring metacognition. Nature Communications, 16(1), 701. *https://doi.org/10.1038/s41467-025-56117-0*

Sandberg, K., Timmermans, B., Overgaard, M., & Cleeremans, A. (2010). Measuring consciousness : Is one measure better than the other? Consciousness and Cognition, 19(4), 1069- 1078. *https://doi.org/10.1016/j.concog.2009.12.013*

Siedlecka, M., Paulewicz, B., & Wierzchoń, M. (2016). But I Was So Sure ! Metacognitive Judgments Are Less Accurate Given Prospectively than Retrospectively. Frontiers in Psychology, 0. *https://doi.org/10.3389/fpsyg.2016.00218*

Wierzchoń, M., Asanowicz, D., Paulewicz, B., & Cleeremans, A. (2012). Subjective measures of consciousness in artificial grammar learning task. Consciousness and cognition, 21(3), 1141-1153. *https://doi.org/10.1016/j.concog.2012.05.012*

https://doi.org/10.7554/eLife.110765.1.sa0