Cognition 254 (2025) 105970

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Anticipating multisensory environments: Evidence for a supra-modal predictive system

Marc Sabio-Albert [a,b,*], Lluís Fuentemilla [a,b,c], Alexis Pérez-Bellido [d]

[a] Department of Cognition, Development and Education Psychology, Faculty of Psychology, University of Barcelona, Spain
[b] Institute of Neuroscience (UBNeuro), University of Barcelona, Spain
[c] Bellvitge Institute for Biomedical Research, Hospitalet de Llobregat, Spain
[d] Department of Basic, Development and Education Psychology, Faculty of Psychology, Autonomous University of Barcelona, Spain

ARTICLE INFO

Keywords:
Prediction
Statistical learning
Implicit learning
Attention
Multisensory
Explicit awareness

ABSTRACT

Our perceptual experience is generally framed in multisensory environments abundant in predictive information. Previous research on statistical learning has shown that humans can learn regularities in different sensory modalities in parallel, but it has not yet determined whether multisensory predictions are generated through a modality-specific predictive mechanism or instead, rely on a supra-modal predictive system. Here, across two experiments, we tested these hypotheses by presenting participants with concurrent pairs of predictable auditory and visual low-level stimuli (i.e., tones and gratings). In different experimental blocks, participants had to attend the stimuli in one modality while ignoring stimuli from the other sensory modality (distractors), and perform a perceptual discrimination task on the second stimulus of the attended modality (targets). Orthogonal to the task goal, both the attended and unattended pairs followed transitional probabilities, so targets and distractors could be expected or unexpected. We found that participants performed better for expected compared to unexpected targets. This effect generalized to the distractors but only when relevant targets were expected. Such interactive effects suggest that predictions may be gated by a supra-modal system with shared resources across sensory modalities that are distributed according to their respective behavioural relevance.

1. Introduction

In recent decades, there has been increasing interest in understanding how predictive information guides and facilitates perception (de Lange et al., 2018; Oliva & Torralba, 2007). Most research on prediction has used experimental paradigms wherein unimodal or cross-modal associations enable the prediction of a single subsequent target (Kok et al., 2012; Kok & Turk-Browne, 2018; Manahova et al., 2018; Meyer & Olson, 2011; Richter et al., 2018; Rosenthal et al., 2018). This approach contrasts with real-life scenarios in which we simultaneously receive predictive sensory inputs from different modalities in parallel.

Previous research on statistical learning (SL), a cognitive process closely tied to prediction (Dale et al., 2012; Turk-Browne et al., 2010) that allows humans to extract regularities from the environment through repeated exposure, has shown that humans can learn statistical regularities from multiple sensory inputs in parallel. For instance, Conway and Christiansen (2006) showed that artificial grammars instantiated in different sensory modalities (visual vs. auditory) could be learned simultaneously. Subsequent studies have replicated similar results (Mitchel & Weiss, 2011; Seitz et al., 2007), providing additional empirical evidence that SL for concurrent streams can take place without interference across sensory modalities. Yet, two unresolved issues remain regarding our inquiry on predictive processing of multisensory environments.

First, the SL studies mentioned above demonstrated that participants had learned the implicit regularities by testing them after the exposure phase, but this does not imply that they used this knowledge throughout the experiment to effectively anticipate inputs. Prediction during SL tasks can only be inferred using trial by trial measures of how the predictability of stimuli affects their processing (Batterink et al., 2015; Dale et al., 2012; Kim et al., 2009; Richter & de Lange, 2019). Therefore, based on previous SL studies it remains unclear whether human observers were actively anticipating multiple sensory inputs in parallel.

Secondly, it is unknown whether predictions across different sensory modalities are completely independent between each other. Conway and Christiansen (2006) showed that once SL was acquired in one

* Corresponding author at: Faculty of Psychology (UB), Pg. de la Vall d'Hebron, 171, Horta-Guinardó, 08035 Barcelona, Spain.
E-mail addresses: msabioal@ub.edu (M. Sabio-Albert), alexis.perez@uab.cat (A. Pérez-Bellido).

https://doi.org/10.1016/j.cognition.2024.105970
Received 6 May 2024; Received in revised form 25 September 2024; Accepted 26 September 2024
Available online 4 October 2024

sensory modality, this knowledge could not be transferred to another sensory modality. Others have also outlined how SL exhibits different characteristics depending on the input's sensory modality (Conway & Christiansen, 2005; Emberson et al., 2011). These modality-specific effects led to the notion that SL may operate independently for each sensory modality (Frost et al., 2015). However, local computations at sensory cortices can be modulated by supra-modal areas such as the hippocampus or the prefrontal cortex (PFC), regions that without being directly involved in the processing of sensory inputs, play an important role in prediction formation (Aitken & Kok, 2022; Clarke et al., 2022; Hindy et al., 2016; Kok & Turk-Browne, 2018; Turk-Browne et al., 2010; van Kesteren et al., 2012). Therefore, if humans can form multiple simultaneous predictions, a logical follow-up question is whether these predictions unfold independently for each sensory modality or instead, they interact with each other, indicating a joint regulation by a supra-modal brain region.

We explored these questions using an adapted version of the probabilistic cueing paradigm in which participants were presented with concurrent but independent probabilistic cue-target transitions, delivered through two distinct sensory channels (auditory and visual). The participants' task was to make perceptual judgments about predictable targets in one modality while ignoring concurrent predictable distractors in the other modality. This allowed us to test if performance in this task was influenced only by target expectations, or also by distractor expectations, implying simultaneous predictive processing. Additionally, the absence or presence of interactions between the predictions in the two sensory modalities would imply parallel and modality-specific processing of predictions or the involvement of a common, supra-modal mechanism, respectively. We conducted two different experiments. In experiment 1 the participants were explicitly aware of the stimuli associations and had to switch attention between modalities across blocks. In experiment 2 they were agnostic about the associations and had to attend only one sensory modality for the whole experiment, while the other remained ignored.

2. Experiment 1

2.1. Material and methods

2.1.1. Participants

We tested 25 psychology students (21 women, 18 right-handed) with normal or corrected to normal vision and hearing, that ranged in age from 18 to 41 years (M = 25 years) belonging to the faculty of Psychology of the University of Barcelona. Participants were compensated with course grades. Since small effect sizes often lack practical significance, the sample size was calculated to detect moderate to large effect sizes ($d = 0.6$) in a repeated measures *t*-test, with a power of 0.8.

All the experimental protocols were approved by the bioethics committee of the same university and were in accordance with the Declaration of Helsinki. All participants received general information about the project and signed an informed consent prior to the performance of the tasks.

2.1.2. Stimuli

The stimuli were generated using PsychoPy version 2021 (Peirce et al., 2019). Visual stimuli consisted of Gabor gratings with 0.7 cycles per degree, that occupied 10 degrees of visual angle and were presented at the center of a grey background. Auditory pure tones were generated online and presented through headphones at a sample rate of 44.1 kHz.

At every trial, a Gabor grating and an auditory tone were simultaneously presented, followed by the simultaneous presentation of a different grating and a different auditory tone. Leading gratings could be vertically (0°) or horizontally (90°) oriented, and trailing gratings could be oriented clockwise (45°) or counterclockwise (135°). Leading auditory tones could have a frequency of 1000 Hz or 1600 Hz, whereas trailing ones had frequencies of either 100 Hz or 160 Hz. Within each sensory modality, leading stimuli acted as predictive cues for the trailing ones. That is, the orientation (frequency) of the leading grating (tone) predicted with a 75 % probability the orientation (frequency) of the trailing grating (tone). These transitional probabilities were defined by transition matrices such as the ones shown in Fig. 1A. For each participant, we balanced the contingencies between each auditory and visual pair. This orthogonalization of the two modalities prevented any possible predictive association between visual and auditory stimuli (Turk-Browne et al., 2008).

2.1.3. Procedure

The experiment started with instructions for the explicit learning phase. Participants' task was to learn the probabilistic associations between leading and trailing stimuli (e.g.: whether a vertical orientation followed by a clockwise orientation was a frequent or infrequent transition). The learning phase spanned four blocks of 40 trials. At the beginning of each block, they were instructed to respond based on the grating pairs (visual blocks) or the pure tones pairs (auditory blocks). As a reminder, a visual icon of an eye or a speaker remained at the bottom of the screen during the blocks. This attentional manipulation switched between visual and auditory modalities in an interleaved fashion (Fig. 1B).

Each trial began with a white fixation dot presented during a time interval randomly selected from a uniform distribution between 750 and 1500 msec. After this, the fixation dot remained on the screen and visual and auditory leading stimuli were presented for 500 msec. After both cues disappeared, a 500 msec fixation dot was followed by a 500 msec presentation of the visual and auditory trailing stimuli. Then, a screen with two response options, "frequent" and "infrequent", and their corresponding keys (z or m, randomized at each block) appeared until the participant gave a response. The fixation dot turned green following correct responses, and red after incorrect ones. There was a third response option, "weak" (spacebar key), to indicate catch trials. In these trials, one of the two trailing stimuli was presented with lower contrast (in the case of gratings) or lower volume (in the case of pure tones). Visual catch trials were included to ensure that participants watched the screen during auditory blocks, but we also included auditory catch trials to equate endogenous attention towards either modality potentially induced by this task instruction. Thus, each block contained eight catch trials (four visual and four auditory).

After the explicit learning phase, participants received instructions for the implicit test phase, which was the focus of our analyses, and it spanned 10 blocks of 72 trials each (5 visual blocks and 5 auditory blocks, interleaved) (Fig. 1B). The stimuli and trial sequences were identical to the previous phase, with the exception that in half of the trials the trailing stimulus of the attended modality (target) had a small deviation (deviant) from its standard orientation (45° or 135°) or frequency (100 Hz or 160 Hz). Participants were instructed to discriminate deviant from standard target stimuli, which made the learned probabilistic associations no longer related to the participants' task. Stimuli presentations, however, continued to follow the learned transitional probabilities.

To balance task difficulty across participants and sensory modalities, the amount of deviation of the targets was controlled with two independent staircase procedures, one for visual blocks and one for auditory blocks. The staircases followed the "3 down 1 up" rule (decrease after three consecutive detections and increase after every failure) targeting a 79.4 % of deviant detections. The step size of deviation changes started from 20° and 20 Hz for visual and auditory targets, respectively; However, this amount dynamically decreased during the experiment based on the number of inversions (change in direction from an increase to a decrease, or vice versa). This was done to efficiently adjust the threshold while decreasing oscillations around it. We also introduced 8 catch trials (4 visual, 4 auditory) in each of the implicit test phase blocks, as well as visual feedback after every response.

The experiment ended with two blocks of a brief explicit recall phase.

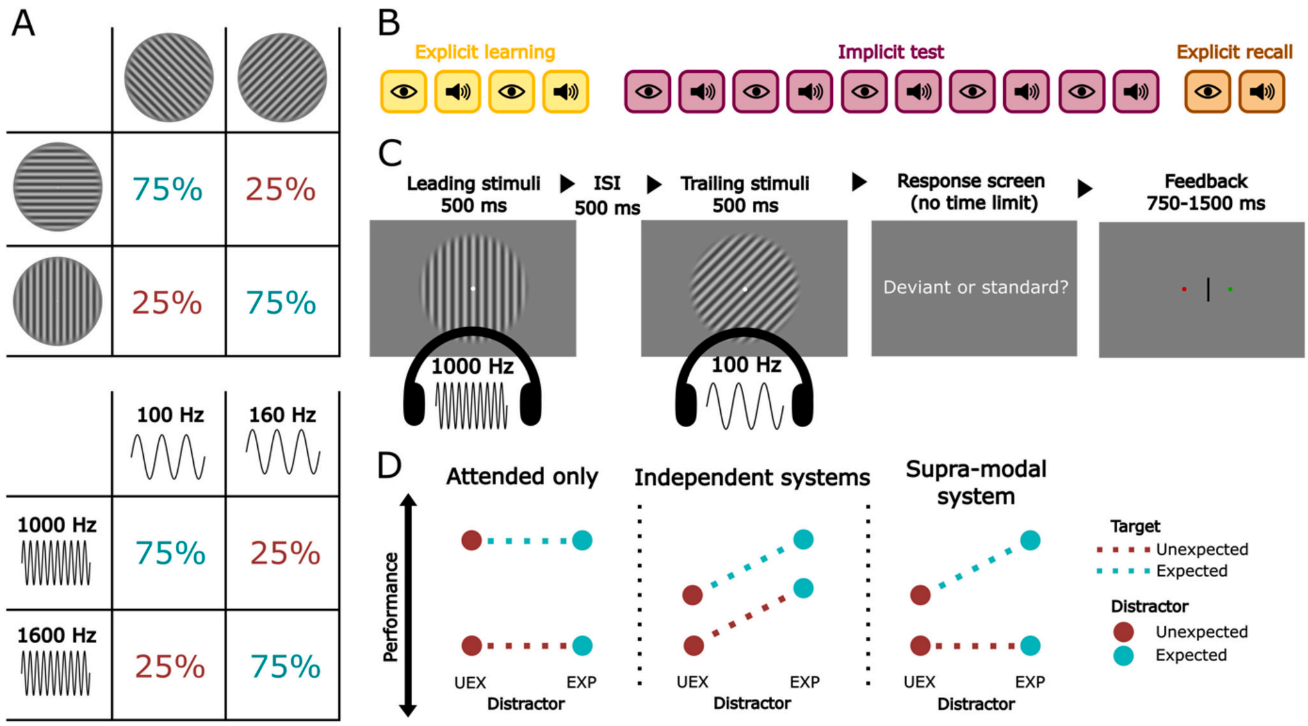

Fig. 1. A) Transition matrices defining the visual (left) and auditory (right) stimuli pair associations. Rows indicate the leading stimuli and columns the trailing stimuli. B) Phases of the experiment. The explicit learning phase (yellow) comprised four blocks of 40 trials (~10 min). The implicit test phase (purple) comprised 10 blocks of 72 trials (~40 min). The explicit recall phase (orange) comprised 2 blocks of 8 trials (~2 min). C) Trial sequence. Every trial started with the concurrent presentation of a leading auditory and visual stimuli, which lasted for 500 ms. After a 500 ms inter stimulus interval, the trailing auditory and visual stimuli were presented, also for 500 ms. Immediately after they disappeared, a response screen with the different response options appeared. There was no time limit for responses. After a response was given, the fixation dot reappeared in green or red, providing feedback. The time interval for this was randomly chosen from 750 ms to 1500 ms. D) Predicted outcomes about expectation effects. Left: if only attended stimuli are predicted, expectations in the unattended modality should have no effect on performance. Center: if attended and unattended predictions are computed independently, they should have additive effects. Right: if attended and unattended predictions are regulated by a unitary system, they should have interactive effects (e.g.: effects of unattended expectations depend on expectation in the attended modality). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

These blocks were unimodal and were used to test if participants retained knowledge about the pairs from the initial learning phase. We presented each of the four possible pairs within a sensory modality twice (8 trials per modality). Participants had to remember if the pair was "frequent" or "infrequent". This phase did not include catch trials and participants did not receive feedback.

2.1.4. Data analyses

We did not exclude any participants based on task performance, as the staircase procedure ensured that all participants had a proportion of correct responses close to 80 % (across participants: M = 83.5 %, std. = 4.8 %). We did however filter out trials in which reaction times exceeded 3 standard deviations from the participant's average. Moreover, we excluded auditory blocks in which no visual catch trial was detected, as we could not guarantee that the participant was watching the visual stimuli. We did not exclude visual blocks in which participants failed to detect auditory catch trials instead, as they received auditory stimulation through headphones. This led to the exclusion of all five auditory blocks of ten participants, two auditory blocks from another participant, and one block from two other participants.

2.1.4.1. Effect of predictions on participants performance. We used one-way repeated measures ANOVA to compare the proportion of correct responses and reaction times between experimental conditions. Our main variables of interest were the expectation of the target (expected vs unexpected transitions in the attended modality) and the expectation of the distractor (expected vs unexpected transitions in the unattended modality). Because we were not interested in modality-specific effects, the main analyses were conducted on the auditory and the visual sensory modalities together, focusing on whether the stimuli were expected/unexpected and attended/unattended. However, for completeness, we replicated the same analyses for each sensory modality separately (see the supplementary materials; Fig. S1 and S2).

2.1.4.2. Expectation effects on sensitivity and response bias. In a perceptual discrimination task, a difference in accuracy between conditions can be explained either by changes in sensitivity (amount of sensory evidence necessary to produce a change in the response), response biases (a general tendency to choose one response over the other) or a combination of both. To find out whether our expectation manipulations influenced participants' sensitivity, response bias or both, we fitted a generalized linear mixed model (GLMM) of the binomial family (logit link function) using pymer4 (Jolly, 2018) to predict participants' responses ("standard" or "deviant"). To estimate participants' sensitivity in discriminating standard from deviant stimuli, we used the absolute sensory difference between standard and presented target stimuli in each trial (difference in orientation ^{o} or frequency Hz, rescaled from 0 to 1 within participant and modality). These trial-by-trial values, that result from the staircase procedure, should be positively correlated with the proportion of "deviant" responses. We included a categorical regressor corresponding to expectation and tested its interaction with the sensory difference. This allowed comparisons between the models' coefficients and derived psychometric curves for the two levels of expectation. A steeper slope of a psychometric curve can be interpreted as an increase in sensitivity (less sensory difference is needed for the observer to discriminate a deviant and respond accordingly). On the other hand, an overall higher or lower intercept would reflect a response bias (a tendency to respond "deviant" or "standard" orthogonal to the sensory difference of the actual targets). Note that while the psychometric curves illustrate the probabilities of a correct answer, our statistical analyses were conducted on the log-odds, which, in simple terms, involve converting a probability-based model into a likelihood-based model. Log-odds follow a positive monotonic relationship with probability, with a log-odds value of 0 corresponding to a probability of 0.5. Log-odds values can easily be converted to probabilities using the inverse logit function. We separately modelled the effects of target and distractor expectations. The models incorporate a random intercept for participant, accounting for each individual's response bias, and a random slope for the effect of sensory difference for each participant, accounting for individual variability in sensitivity across our sample.

2.1.4.3. Modelling interaction between attended and unattended modality expectations. We next aimed to examine whether the effects of target and distractor expectations were additive or interactive (Fig. 1D). The rationale behind this is that if two parallel processes are influenced separately by two external factors (in our study, the expectation of visual and auditory stimuli), the joint impact of these two factors on performance will be additive. Alternatively, an interaction between them (i.e.: the effect of one factor differs depending on the levels of the other factor) would indicate a dependence between both processes and potentially shared processing (Sternberg, 1969).

The first alternative would imply that predictions in each sensory modality are independent from each other, and possibly orchestrated by independent systems. The second alternative would imply that predictions in the different modalities are interdependent, and therefore must be jointly regulated by a supra-modal system that supervises information predictability across sensory modalities. To arbitrate between these two possibilities, we fitted a GLMM that predicted the probability of a correct response as a function of target and distractor expectations, as well as their interaction. We included the attended modality (visual or auditory) as a third regressor, as well as its interaction with the other two regressors, improving the model's fit and accounting for potential modality-specific effects. We included a random intercept for each subject accounting for individual differences in baseline accuracy.

2.1.4.4. Assessing the relationship between expectations and explicit knowledge. To determine whether the explicit knowledge about stimuli associations modulated the effects of sensory predictions, we classified each stimulus pair as "recalled" or "not recalled" for each participant. A stimulus pair was classified as "recalled" if they were correctly reported as "frequent" or "infrequent" on at least 3 of the 4 explicit recall phase trials in which it was presented (e.g.: in the case of a participant for whom vertically oriented gratings had been followed by clock-wise gratings on 75 % of trials, a correct response would either be "frequent" to vertical followed by a clock-wise grating or "infrequent" to vertical followed by a counter-clockwise grating). Then we fitted two logistic GLMMs, one for targets and another for distractors, predicting the probability of being correct as a function of expectation, recall and their interaction. We also included an interaction with the attended modality, improving the model's fit. The models included a random intercept per participant, accounting for participant's baseline accuracy.

2.2. Results

Very few trials (1.9 %) were excluded due to outlier RTs further than 3 SDs from the participant mean. The analyses of performance in the learning phase indicated that participants acquired explicit knowledge about the probabilistic transitions. Their accuracy was higher in the second compared to the first learning blocks of each modality meaning that their performance improved with practice (visual accuracy in block 1 = 66.3 % and in block 2 = 80 %, $t(24) = -6.66$, $p < .0001$, $d = -1.16$; auditory accuracy in block 1 = 66.5 % and in block 2 = 74.62 %, $t(24) = -3.67$, $p = .001$, $d = -0.61$). This accuracy improvement was accompanied by faster reaction times (RTs; visual RTs in block 1 = 1145 msec and in block 2 = 782 msec, $t(24) = 5.07$, $p < .001$, $d = 1.15$; auditory RTs in block 1 = 1266 msec and in block 2 = 788 msec, $t(24) = 4.71$, $p < .001$, $d = 1.03$).

During the subsequent implicit test phase, the proportion of correct responses in both modalities did not differ from the 80 % adaptive staircase target threshold (visual blocks: 0.86, 95 % *CI*[0.79, 0.9]; auditory blocks: 0.78, 95 % *CI*[0.68, 0.83]). In subsequent analyses we analyzed the visual and auditory conditions together, but we included attended sensory modality as an independent variable to control for potential modality-specific effects.

2.2.1. Participants concurrently predicted targets and distractors

To assess whether participants concurrently predicted trailing stimuli of both sensory modalities, we tested if their performance in the implicit test phase depended on the expectations for both targets and distractors.

For the attended targets, participants had a higher accuracy in expected (84.1 %) compared to unexpected target trials (78.4 %; $F(1,24) = 11.68$, $p = .002$, *partial* $\eta 2 = 0.33$; Fig. 1A), and responded faster in expected (567 msec) than in unexpected target trials (593 msec; $F(1,24) = 5.06$, $p = .033$, *partial* $\eta 2 = 0.17$). Crucially, participants were also more accurate (83.5 % vs 80.03 %; $F(1,24) = 13.19$, $p = .001$, *partial* $\eta 2 = 0.35$) and faster (567 msec vs 595 msec; $F(1,24) = 11.41$, $p = .002$, *partial* $\eta 2 = 0.32$) when distractors in the unattended modality were expected than when they were unexpected (Fig. 1B). This result indicates that the participants kept track of the two probabilistic structures simultaneously.

From these results it is unclear how specifically distractors influenced performance. In the following analyses, we calculated how distractor expectations affected participants' sensitivity, which in our task relates to the ability to distinguish a deviant from a standard target stimulus. We also examined response bias, such as an increased tendency to report "deviant" after an expectation violation.

2.2.2. Predictions about targets and distractors affected sensitivity and bias

We tested whether expectations about targets and distractors could be differentiated in terms of sensitivity and response bias. We fitted two GLMMs predicting response ("deviant" or "standard") as a function of the sensory difference between presented targets and standard stimulus. The first model incorporates a regressor for the expectation of the target stimulus, and the second one for the expectation of the distractor (see Data analysis section). The GLMM for the attended modality revealed an overall lower probability to report expected targets as deviants, indicated by the expectation coefficient ($Z = -8.84$, $p < .001$), and a significantly higher sensitivity for expected vs. unexpected targets, as indicated by the interaction between sensory difference and expectation ($Z = 5.04$, $p < .001$) (Fig. 2C).

The GLMM for the unattended modality revealed a similar pattern of results; a main effect of distractor expectation ($Z = -3.97$, $p < .001$) and a significant interaction between the expectation and sensory difference of distractors ($Z = -2.02$, $p = .044$; Fig. 2D) indicated that participants had a lower response bias and higher sensitivity when distractors were expected.

Therefore, our results show that target and distractor predictability influenced both sensitivity and response bias.

2.2.3. Effects of expectation interacted across modalities

Next, we tested whether the effect of attended and unattended predictions interacted. We fitted a GLMM predicting correct responses as a function of target and distractor expectations (see Data analyses). The interaction between these two regressors was significant ($Z = 2.18$, $p = .029$). Post hoc comparisons between expectation estimates revealed that distractors predictability had a significant effect on participants

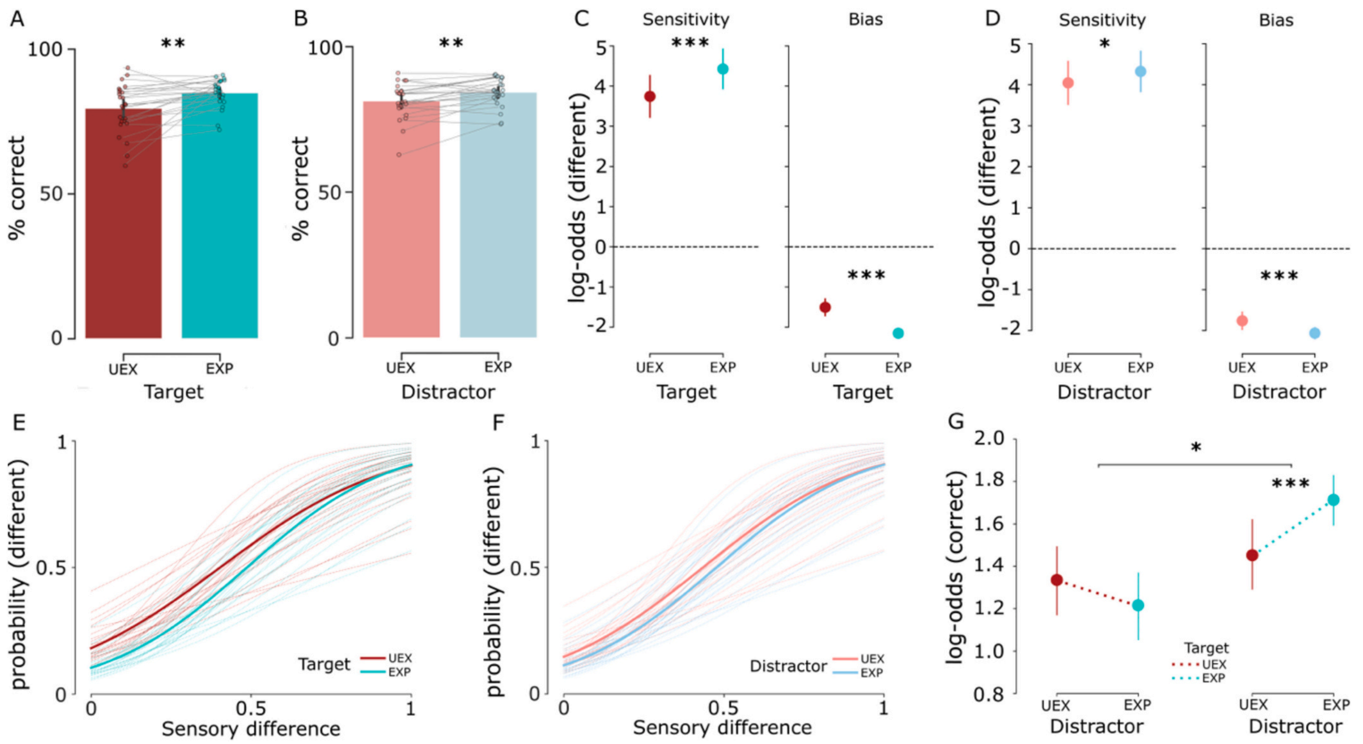

Fig. 2. A) Comparison of accuracy in unexpected target trials (red) and expected target trials (blue). Dots correspond to the accuracy of individual participants, connected by lines across conditions. B) Same as panel A but comparing unexpected distractor trials (light red) with expected distractor trials (light blue). C) GLMM parameter estimates for the attended modality and their 95 % confidence intervals. The column on the left contains the estimated coefficient for sensory difference for unexpected target trials (red) and expected target trials (blue). The column on the right shows the estimated coefficient of unexpected targets (red) and expected targets (blue). D) GLMM parameter estimates for the unattended modality and their 95 % confidence intervals. E) GLMM predicted logistic regression functions for the attended modality. Solid lines are for the fixed effects and dashed lines for individual participant random effects. Red lines are for unexpected targets, and blue lines for expected targets. F) GLMM predicted logistic regression function for the unattended modality. G) Parameter estimates of the interaction GLMM. Left column: distractor expectations when targets were unexpected. Right column: distractor expectations when targets were expected. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

accuracy when targets were expected ($Z = -3.43$, $p < .001$) but not when they were unexpected ($Z = 1.32$, $p = .18$). The lack of difference between the two conditions was not caused by a floor effect in unexpected target trials (accuracy = 78.4 %). This result demonstrates that participants were only sensitive to predictions in the unattended modality when targets were expected. This dependence of unattended prediction on the attended predictions' fulfillment suggests that a common system regulated visual and auditory expectations, contrasting with the alternative possibility that concurrent predictions in different sensory modalities are processed independently.

2.2.4. *Explicit knowledge only affects predictive processing of attended stimuli*

At the end of the experiment, during the "explicit recall" phase, participants correctly classified 66.25 % of presented pairs (60.5 % for auditory pairs and 72 % for visual pairs) (Fig. 3A). This result indicates

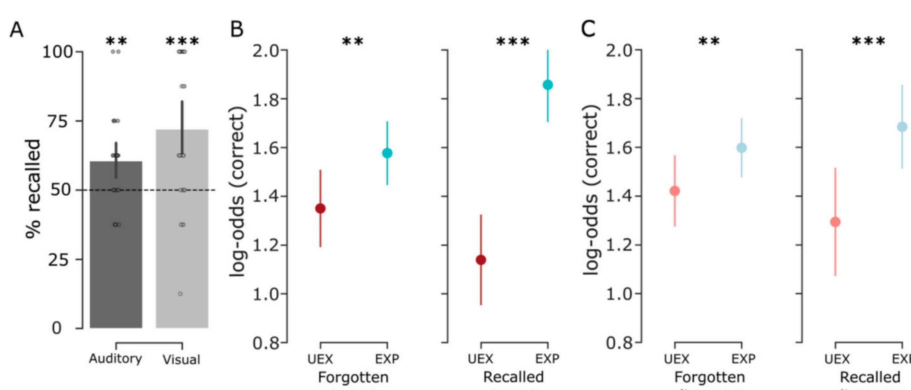

Fig. 3. A) Percentage of correct recalls of auditory and visual pairs in the explicit recall phase. Individual participants are represented by dots. The horizontal line at 50 % is the chance level performance. B) Parameter estimates and confidence intervals of the GLMM for the attended modality. Left column: Estimated log-odds of a correct response as a function of target expectation when the attended pair was forgotten in the explicit recall phase. Right column: same as left column but when the attended pair was recalled in the explicit recall phase. B) Estimates and confidence intervals of the GLMM for unattended modality. As in fig. A, the estimated log-odds of correct response as a function of distractor expectation (colour) and if the unattended modality pair was forgotten or recalled in the explicit recall phase.

that they retained explicit knowledge throughout the experiment about the visual (binomial test, $p < .001$) and auditory stimuli associations ($p = .004$).

To test whether expectation effects depended on the explicit knowledge of the probabilistic associations, we then fitted two GLMMs, separately for attended and unattended modalities, that predicted correct responses in the implicit test phase as a function of stimuli expectations, of whether stimuli pairs were explicitly recalled or not, and their interaction (see Data analyses section). The model fitted to the attended stimuli yielded a significant interaction between expectation and explicit recall ($Z = 2.74$, $p = .006$), indicating a larger effect of target expectations when the presented pair was later recalled ($Z = -7.65$, $p < .001$) than when it was forgotten ($Z = -3.39$, $p = .001$) (Fig. 3B). In contrast, this interaction was not significant in the model for the unattended distractors ($Z = 1.47$, $p = .14$; distractor recalled: $Z = -3.45$, $p = .001$; distractor forgotten: $Z = -2.82$, $p = .005$) (Fig. 3C), demonstrating that explicit knowledge about the associations only contributed to the predictions when the stimuli were attended. In summary, our results demonstrate that participants are able to predict targets and distractors despite not being explicitly aware about their probabilistic associations, confirming that implicit mechanisms of associations underlie our findings. Additionally, our results also indicate that predictions for attended stimuli are subject to additional top-down controlled cognitive processes that depend on the explicit awareness of the leading-trailing stimuli associations.

2.3. Discussion

We replicated previous studies (Kok et al., 2012; van Ede et al., 2012) by demonstrating that probabilistic cues, which were not essential for the task, could either enhance or hinder participants' behavioural performance. Remarkably, these effects were caused by the expectations about concurrently presented targets and distractors in different sensory modalities, implying that participants simultaneously predicted both stimuli. The fact that expectations about distractors only influenced performance when targets were expected suggests the existence of a shared predictive system overarching the different modalities. In addition, we found that predictions affected performance even when participants were not explicitly aware of their associations, highlighting the automatic nature of these predictive processes. Nevertheless, retaining explicit knowledge of the associations boosted the effects of prediction for attended targets, pointing towards additional processing gated by attention that relies on explicit knowledge.

Based on how participants performance was affected by task-irrelevant stimuli, we can indirectly conclude that participants monitored prediction fulfillment and violations even for unattended events. However, given participants' awareness of the associations and that they had to switch attention between modalities in different blocks, they might have felt inclined to verify whether the contingencies remained stable in the unattended modality rendering it not completely unattended. To rule out this hypothesis, and to ensure that the unattended modality is completely irrelevant during the whole experiment, we conducted a second experiment in which participants did not have explicit knowledge about the existence of associations between the stimuli in any modality. In addition, one of the modalities was consistently maintained as unattended throughout the experiment.

3. Experiment 2

This experiment investigated whether the prediction effects observed for unattended stimuli in experiment 1 could be caused by participants' explicit knowledge of the underlying probabilistic structure for both attended and unattended stimuli. To explore this, in this second experiment participants performed the same task as in experiment 1, but we did not train them to learn the associations, nor informed them about their existence. Additionally, we ensured that the unattended modality remained behaviourally irrelevant during the whole task. To achieve this, we tested two distinct groups of participants: one group discriminated visual stimuli with auditory distractors, and another group discriminated auditory stimuli with visual distractors.

3.1. Material and methods

3.1.1. Participants

We tested 65 psychology students with normal or corrected to normal vision and hearing in the faculty of Psychology of the University of Barcelona. 34 were assigned to the auditory group (i.e. attended to auditory stimuli), and 31 to the visual group (i.e. attended to visual stimuli). Participants were compensated with course grades. We did not exclude any participant from the visual group. From the auditory group we excluded four participants who did not detect any visual catch. Following this criterion, we excluded 9 blocks from another 3 participants, 8 blocks form 1 participant, 5 blocks from 1 participant, and 2 blocks from 3 participants. We excluded 4 more participants for not getting 60 % of correct responses. Finally, for every participant, we filtered out trials exceeding 3 standard deviations from their mean reaction time (1.98 % of trials). Thus, our final sample consisted of 57 participants, 26 in the auditory group (23 women, 21 right-handed, 18–24 years old, M = 19.4) and 31 in the visual group (26 women, 29 right-handed, 18–47 years old, M = 22.5). The final sample sizes were sufficient to detect moderate to large effects ($d = 0.6$) with a power level of 0.8 within each group.

3.1.2. Stimuli

The stimuli were the same as in the first experiment, and participants did the task in the same room and under the same experimental conditions.

3.1.3. Procedure

Participants were randomly assigned to the auditory or visual group. Because we wanted to test whether participants could implicitly learn the probabilistic associations between leading and trailing stimuli, they did not receive information about the pair associations, and we substituted the explicit learning phase of experiment 1 for a familiarization phase. This phase had the same duration as the learning phase in the first experiment (4 blocks of 40 trials), but participants were only instructed to detect catch stimuli of either modality, which appeared on 8 trials (4 visual and 4 auditory). The implicit test phase was identical to the one in the first experiment, except that now participants had to attend to the same modality in the 10 blocks. They also had to respond to catch trials from both modalities. After this phase there was another one analogous to the explicit recall phase in experiment 1, although we first had to inform them about the transitional probabilities. Then, following the same procedure as in the first experiment, they had to indicate whether presented pairs had appeared frequently or infrequently during the implicit test phase.

3.2. Results

We analyzed the data from the visual and auditory groups together, only focusing on the effects of expectations in the attended and unattended modalities. However, to ensure that the pattern of results was not mainly driven by one of the experimental groups, we replicated the analyses separately for each attended modality (supplementary materials S3 and S4). In both groups, the proportion of correct responses did not differ significantly from the expected 80 % of the staircase procedure (visual group: 0.87, 95 % *CI*[0.76, 0.92]; auditory group: 0.75, 95 % *CI* [0.62, 0.86]).

3.2.1. Concurrent predictions for targets and distractors

To test if participants formed predictions about targets and distractors during the implicit test phase, we first contrasted accuracies (proportion of correct responses) and RTs between expected vs. unexpected targets, as well as expected vs. unexpected distractors. We replicated the results of experiment 1. Participants had a better accuracy to expected (83.1 %) than to unexpected targets (77.8 %; $F(1,54) = 61$, $p < .001$, *partial* $\eta 2 = 0.53$; Fig. 4A), and also responded faster to them (536 msec. vs. 562 msec; $F(1,54) = 19.75$, $p < .001$, *partial* $\eta 2 = 0.27$). Likewise, when distractors were expected the accuracy was higher (82.8 % vs. 78.9 %; $F(1,54) = 26.56$, $p < .001$, *partial* $\eta 2 = 0.33$) (Fig. 4B) and the reaction times faster (540 msec. vs. 553 msec, $F(1,54) = 5.74$, $p = .02$, *partial* $\eta 2 = 0.09$) than when they were unexpected. These results suggest that participants could form predictions about task-irrelevant stimuli, without being explicitly aware about their probabilistic structure. Moreover, the observed effects of predictions about unattended distractors highlight the automaticity of predictive processing.

3.2.2. *Attended and unattended predictions change sensitivity and response bias*

We explored whether prediction related changes in accuracy were explained by changes in the sensitivity and/or response biases. As in experiment 1, we fitted GLMMs predicting responses as a function of the sensory difference between standard and deviant targets (Data analysis section), as well as attended modality and unattended modality expectations (in two separate models). Participants had a lower tendency to respond "different" when targets were expected, indicated by the main effect of target expectation ($Z = -8.78$, $p < .001$) (Fig. 4C). Expected targets were also associated to a higher sensitivity than unexpected targets, indicated by the significant interaction between sensory difference and target expectation ($Z = 6.57$, $p < .001$) (Fig. 4C). The GLMM for the unattended modality yielded analogous estimates: a reduced bias to respond "different" for expected distractors ($Z = -3.44$, $p < .001$), and a significant interaction between sensory difference and distractor expectations ($Z = 2.47$, $p = .013$; Fig. 4D). Thus, our results replicate experiment 1 results by showing that expected targets and distractors improve participants' sensitivity while also reducing their response bias to report "different", compared to unexpected targets and distractors.

3.2.3. *Distractor predictions only affected performance when targets were expected*

Next, we sought to determine if the interaction between expectations in the attended and unattended modalities that we observed in experiment 1 persisted when the distractors were completely irrelevant for the task. We fitted a GLMM that predicted the probability of correct responses as a function of target and distractor expectations and their interaction. The interaction between target and distractor predictions was marginally significant ($Z = 1.92$, $p = .054$). Nevertheless, we replicated the pattern of results of experiment 1 (Fig. 4G); when targets were expected, the distractors had a significant effect on accuracy ($Z = -5.03$, $p < .001$), but not when targets were unexpected ($Z = -0.82$, $p = .42$). This result is consistent with our previous finding that predictions across sensory modalities are not independent.

3.2.4. *Predictions without explicit awareness*

Participants did not correctly classify pairs of the ignored modality above chance level in the explicit awareness phase (Fig. 5A; visual group: 53.6 % of correct classification, binomial test $p = .28$; auditory group: 47.5 % of correct classification, binomial test $p = .52$).

The visual group did not classify the visual attended pairs above chance either (52.4 %, $p = .48$). In contrast, the auditory group correctly classified 58.5 % of pairs, reaching a significantly above-chance performance ($p = .001$). This means that some participants in the auditory group may have become aware of the pair associations during the experiment. To test if there was any interaction between predictions and

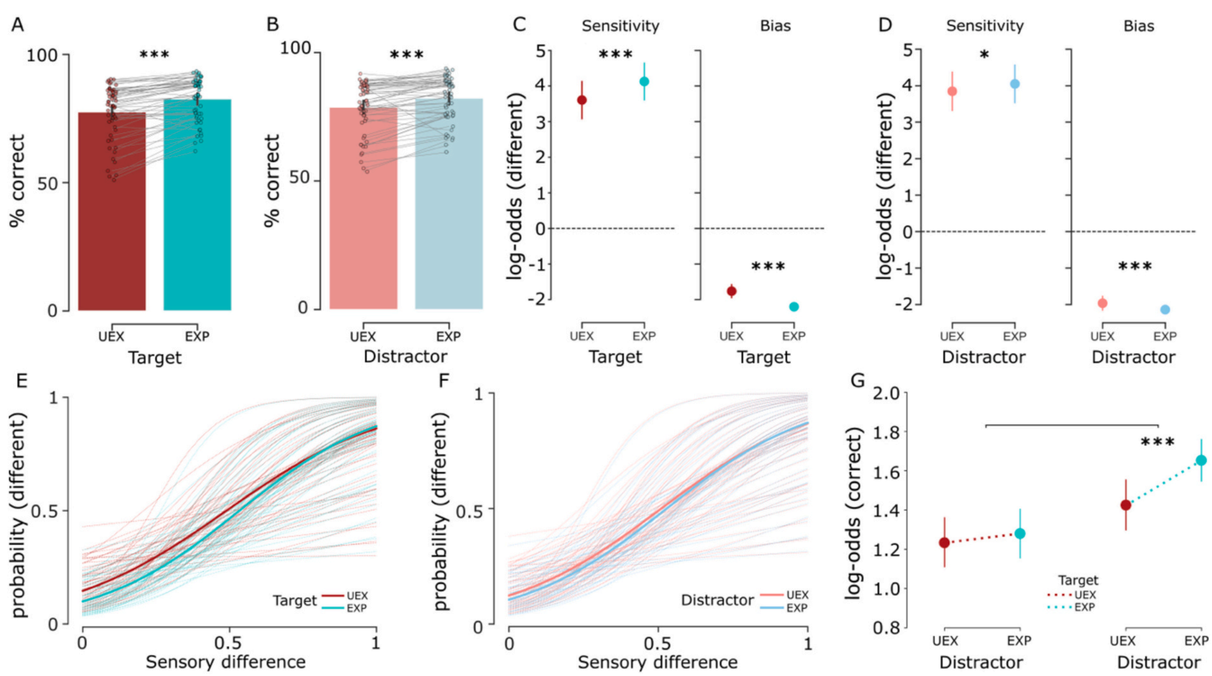

Fig. 4. A) Comparison of accuracy in unexpected target trials (red) and expected target trials (blue). Dots correspond to the accuracy of individual participants, connected by lines across conditions. B) Same as panel A but comparing unexpected distractor trials (light red) with expected distractor trials (light blue). C) Parameter estimates of the GLMM for the attended modality and their 95 % confidence intervals. The column on the left contains the parameter estimates for sensory difference for unexpected target trials (red) and expected target trials (blue). The column on the rate shows the parameter estimates of unexpected targets (red) and expected targets(blue). D) Parameter estimates of the GLMM for the unattended modality and their 95 % confidence intervals. E) GLMM predicted logistic regression functions for the attended modality. Solid lines are for the fixed effects and dashed lines for individual participant random effects. Red lines are for unexpected targets, and blue lines for expected targets. F) GLMM predicted logistic regression functions for the attended modality. G) Estimates of the interaction GLMM. Left column: distractor expectations when targets were unexpected. Right column: distractor expectations when targets were expected. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

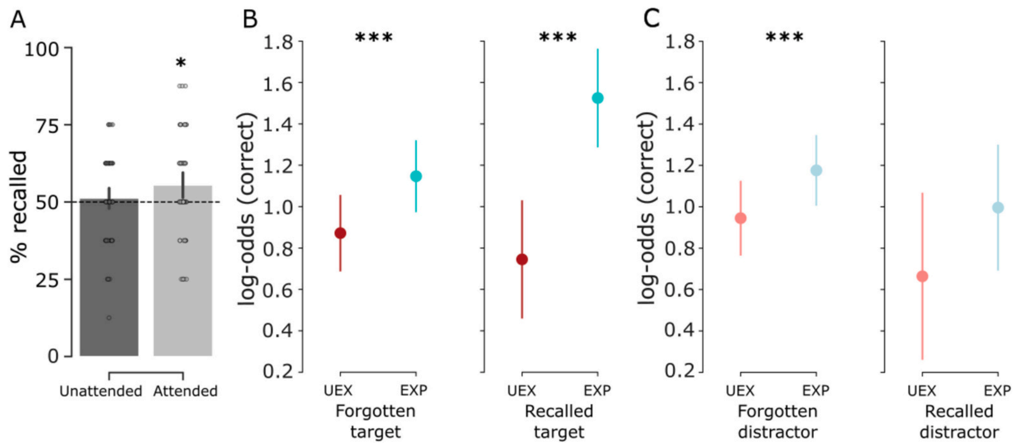

Fig. 5. A) Percentage of correct recalls of attended and unattended modality pairs in the explicit phase. Individual participants are represented by dots. The horizontal line at 50 % is the chance level performance. B) Estimates and confidence intervals of the GLMM for the attended auditory modality. Left column: Estimated log-odds of a correct response in the implicit test phase as a function of target expectation when the participants were not explicitly aware of the auditory pair. Right column: same as left column but when they were explicitly aware of the auditory pair. B) Estimates and confidence intervals of the GLMM for the unattended visual modality. As in fig. A, the estimated log-odds of correct response as a function of distractor expectation (colour) and if the participant was explicitly aware of the visual modality pair or not (column).

explicit awareness during the implicit phase of experiment 2, we fitted a GLMM with data of the auditory group only, predicting the probability of correct responses as a function of auditory target expectation, explicit awareness, and their interaction. For completeness, we did the same analyses for visual distractor expectations, despite the lack of evidence of explicit awareness about them.

The model fitted to the auditory group's responses revealed a significant interaction between attended expectations and explicit awareness ($Z = 3.56$, $p < .001$). Similarly to the results of experiment 1, participants in the auditory group showed larger expectation effects when they were explicitly aware of the auditory pair associations ($Z = -5.84$, $p < .001$), than when unaware of them ($Z = -5.72$, $p < .001$) (Fig. 5B). However, we suggest caution when interpreting this result due to the low number of observations, as only 6 participants correctly classified one auditory pair in the explicit phase. In the GLMM with visual distractor expectation, the interaction with explicit awareness was not significant ($Z = 0.49$, $p = .6$) (Fig. 5C). This result was not surprising as participants did not exhibit significant explicit learning in the visual group.

Taken together, these results indicate that the expectation effects observed in experiment 2 largely emerged due to implicit learning mechanisms. However, some participants may have acquired explicit knowledge about attended pairs after repeated exposure which, as in experiment 1, was associated to enhanced predictive processing only for those stimuli.

3.3. Discussion

In this second experiment, we investigated whether the associations between the unattended modality stimuli must be explicitly learned prior to producing the predictive effect. Thus, in this experiment participants were not trained to learn the pair associations, and they had to attend only one sensory modality throughout the entire implicit test phase. This experimental manipulation produced the expected result. Participants did not manifest explicit awareness about the association between the stimuli presented in the unattended modality. Even so, we replicated the results of experiment 1: participants performed better in trials when targets and distractors were expected, demonstrating once again that participants simultaneously kept track of the probabilistic associations in different modalities. We also replicated the interaction by which distractor predictions stopped influencing performance when the targets were unexpected, further supporting our hypothesis that predictions depend on a supra-modal system.

In addition, we found that participants were able to explicitly recognize items at the end of the experiment if they had been attended to. However, this effect was specific to auditory associations, as we could not find evidence that participants in the visual group explicitly learned the attended visual pairs. This asymmetry may be because the auditory modality is typically favored over visual stimulation in sequential statistical learning tasks (Conway & Christiansen, 2005; Emberson et al., 2019; Saffran, 2002). Interestingly, we replicated the effect that participants with larger recognition of the auditory associations during the recall phase also experienced stronger prediction effects.

In summary, results of experiment 2 revealed that the majority of predictive phenomena reported in the present work do not rely on explicit awareness about stimuli associations, but that they are underpinned by neural mechanisms that automatically extract and react to environmental statistical regularities.

4. General discussion

In this study, we showed that human observers automatically track statistical regularities in parallel across the visual and auditory sensory modalities, irrespective of the focus of attention. We derive this conclusion from the observation that performance in the attended modality did not only depend on whether the task-relevant targets were expected or unexpected, but also on the predictability of the distractors in the unattended modality. Specifically, we found that distractors' predictions only affected participants' performance when predictions in the attended modality were fulfilled. We propose that this interaction between attended and unattended expectations can be explained because predictions across different sensory modalities are jointly modulated by a supra-modal system. This system might have limited resources and therefore should distribute them efficiently to handle complex, multielement scenes. For instance, when behaviourally relevant stimuli do not match the predictions, the processing of the unexpected but task-relevant stimulus is prioritized at the expense of detecting violations of less relevant predictable information. In summary, our results provide empirical evidence that human observers are not just capable of concurrently learning statistical regularities instantiated in different sensory modalities (Conway & Christiansen, 2006; Mitchel & Weiss, 2011; Seitz et al., 2007), but also of generating multisensory and independent predictions on a trial-by-trial basis. Importantly, although we combined the two sensory modalities in our main analyses, complementary analyses (see supplementary materials) demonstrated that the effects were symmetrical across sensory modalities. In other words, we found the same pattern of results when participants attended to either the visual or auditory modality. This indicates that this mechanism is not specific to one sense but rather a general principle of perceptual processing.

To understand whether prediction violations affected participants' performance at a perceptual level or induced response biases, we modelled their responses using logistic GLMMs. This analysis revealed a dual effect of predictions: violated expectations about targets and distractors were associated with a lower sensitivity in discriminating deviant stimuli, as well as a higher bias to report targets as deviants. Hence, the modulations in sensitivity suggest that predictions changed the perceptual processing of the stimuli. We should note that based on the behavioural data it is difficult to determine whether predictions enhanced the processing of expected targets or instead impaired the processing of unexpected targets (or both). This is because our experimental design lacks a neutral condition to separately contrast the effects of fulfilled and violated expectations (Feuerriegel et al., 2021). Thus, based on previous studies, the difference in sensitivity between expected and unexpected targets could be related to an enhanced processing of the expected compared to the unexpected stimulus (Cheadle et al., 2015; Pinto et al., 2015; Wyart et al., 2012). However, the explanation for how distractor expectations affect the perceptual sensitivity of the target stimuli is less straightforward. In line with the hypothesis proposed by Alink and Blank (2021), we speculate that the detection of an unexpected event in another modality could transiently orient attention away from targets, impairing their perceptual processing. This explanation is supported by the increased bias to categorize a target as deviant when a distractor is unexpected. Such bias indicates that participants indeed detect and react towards distractor violations, possibly by attributing the sensation of detecting a violation to the discrimination task (which manifested as an increased bias to report targets as deviants). A different, but not mutually exclusive possibility is that expected distractors are easier to ignore, leaving more free resources to process the target stimuli (see Alink and Blank (2021) for an in-depth discussion on the interdependence of attention and expectation). Further research using neuroimaging methods might help to arbitrate between these different hypotheses.

Another key result that we replicated in both experiments is that participants capitalized on explicit knowledge about the pair associations to predict the stimuli. However, this "explicit strategy" only manifested when the modality was attended. Our results support emerging hypotheses considering statistical learning, an ability classically assumed to operate incidentally and implicitly (Fiser & Aslin, 2001; Turk-Browne et al., 2005), to involve "optional" explicit learning mechanisms mediated by attention (Arciuli, 2017; Conway, 2020). Such a double functional dissociation of top-down controlled and implicit mechanisms' contributions to predictions could also explain divergent patterns of results in the prediction literature, like the presence of dampening or sharpening of neural activity patterns elicited by expected stimuli (see Press et al. (2020) for further discussion on this problem). We speculate that top-down modulated predictions may be analogous to endogenous attentional processing, leading to sharpened neural representations arising from the suppression of sensory units that are tuned away from expected inputs. This mechanism would be more evident in probability cueing paradigms in which due to the low number of stimuli contingencies, is easier to assimilate the probabilistic structure of the task explicitly, in consistency with previous studies (Bell et al., 2016; Kok et al., 2012, 2019; Yon et al., 2018). On the other hand, the implicit predictive mechanism may "dampen" the neural representations due to the cancellation of expected information. Such mechanisms would be more evident in purer statistical learning tasks with a larger number of items, as learning of the association would be more difficult to develop during the task (Blank & Davis, 2016; Han et al., 2019; Kumar et al., 2017; Meyer & Olson, 2011; Richter et al., 2022). This hypothesis should be tested in new experiments.

5. Conclusions

In summary, this study demonstrates that human observers can predict inputs in multiple sensory modalities simultaneously. Moreover, these predictions are not entirely independent across sensory modalities, as predictions about distractors only affected participants' performance when predictions in the attended modality were fulfilled. Our results are consistent with the idea that predictions in different sensory modalities are coordinated at a supra-modal level. This represents an optimal strategy, as the processing resources dedicated to each prediction are efficiently prioritized based on their behavioural relevance. Additionally, we showed that the predictive system leverages explicit knowledge about attended stimuli but not unattended stimuli, revealing a dissociable impact of implicit and explicit learning on predictions. Our findings pave the way for more ecologically valid experiments in which the human perceptual system is challenged to predict multiple items simultaneously, potentially revealing additional mechanisms that contribute to the functioning of our predictive system.

PsyArXiv preprint

https://osf.io/preprints/psyarxiv/6nkbp

Data and code available

https://github.com/msabioal7/multiPred

CRediT authorship contribution statement

Marc Sabio-Albert: Writing – original draft, Visualization, Software, Methodology, Investigation, Data curation, Conceptualization. **Lluís Fuentemilla:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization. **Alexis Pérez-Bellido:** Writing – review & editing, Supervision, Project administration, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

None.

Data availability

Data and code is publicly accessible through the following link: https://github.com/researcher000/multiPred

Acknowledgements

This work was supported by the Spanish Ministerio de Ciencia, Innovación y Universidades, which is part of Agencia Estatal de Investigación (AEI), through the projects RTI2018-100977-J-I00 and RYC2022-037652-I to A.P.B. and the projects PID2019-111199GB-I00 and PID2022-140426NB-I00 to L.F. (funded by MCIN/AEI/10.13039/501100011033/ and FEDER a way to make Europe). We thank CERCA Programme/Generalitat de Catalunya for institutional support.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2024.105970.

References

Aitken, F., & Kok, P. (2022). Hippocampal representations switch from errors to predictions during acquisition of predictive associations. *Nature Communications, 13*(1), 3294. https://doi.org/10.1038/s41467-022-31040-w

Alink, A., & Blank, H. (2021). Can expectation suppression be explained by reduced attention to predictable stimuli? *NeuroImage, 231*(February), Article 117824. https://doi.org/10.1016/j.neuroimage.2021.117824

Arciuli, J. (2017). The multi-component nature of statistical learning. *Philosophical Transactions of the Royal Society, B: Biological Sciences, 372*(1711), 20160058. https://doi.org/10.1098/rstb.2016.0058

Batterink, L. J., Reber, P. J., Neville, H. J., & Paller, K. A. (2015). Implicit and explicit contributions to statistical learning. *Journal of Memory and Language, 83*, 62–78. https://doi.org/10.1016/j.jml.2015.04.004

Bell, A. H., Summerfield, C., Morin, E. L., Malecek, N. J., & Ungerleider, L. G. (2016). Encoding of stimulus probability in macaque inferior temporal cortex. *Current Biology, 26*(17), 2280–2290. https://doi.org/10.1016/j.cub.2016.07.007

Blank, H., & Davis, M. H. (2016). Prediction errors but not sharpened signals simulate multivoxel fMRI patterns during speech perception. *PLoS Biology, 14*(11), Article e1002577. https://doi.org/10.1371/journal.pbio.1002577

Cheadle, S., Egner, T., Wyart, V., Wu, C., & Summerfield, C. (2015). Feature expectation heightens visual sensitivity during fine orientation discrimination. *Journal of Vision, 15*(14), 14. https://doi.org/10.1167/15.14.14

Clarke, A., Crivelli-Decker, J., & Ranganath, C. (2022). Contextual expectations shape cortical reinstatement of sensory representations. *Journal of Neuroscience, 42*(30), 5956–5965. https://doi.org/10.1523/JNEUROSCI.2045-21.2022

Conway, C. M. (2020). How does the brain learn environmental structure? Ten core principles for understanding the neurocognitive mechanisms of statistical learning. *Neuroscience & Biobehavioral Reviews, 112*, 279–299. https://doi.org/10.1016/j.neubiorev.2020.01.032

Conway, C. M., & Christiansen, M. H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 31*(1), 24–39. https://doi.org/10.1037/0278-7393.31.1.24

Conway, C. M., & Christiansen, M. H. (2006). Statistical learning within and between modalities: Pitting abstract against stimulus-specific representations. *Psychological Science, 17*(10), 905–912. https://doi.org/10.1111/j.1467-9280.2006.01801.x

Dale, R., Duran, N. D., & Morehead, J. R. (2012). Prediction during statistical learning, and implications for the implicit/explicit divide. *Advances in Cognitive Psychology, 8*(2), 196–209. https://doi.org/10.2478/v10053-008-0115-z

van Ede, F., de Lange, F. P., & Maris, E. (2012). Attentional cues affect accuracy and reaction time via different cognitive and neural processes. *Journal of Neuroscience, 32*(30), 10408–10412. https://doi.org/10.1523/JNEUROSCI.1337-12.2012

Emberson, L. L., Conway, C. M., & Christiansen, M. H. (2011). Timing is everything: Changes in presentation rate have opposite effects on auditory and visual implicit statistical learning. *Quarterly Journal of Experimental Psychology, 64*(5), 1021–1040. https://doi.org/10.1080/17470218.2010.538972

Emberson, L. L., Misyak, J. B., Schwade, J. A., Christiansen, M. H., & Goldstein, M. H. (2019). Comparing statistical learning across perceptual modalities in infancy: An investigation of underlying learning mechanism(s). *Developmental Science, 22*(6), Article e12847. https://doi.org/10.1111/desc.12847

Feuerriegel, D., Vogels, R., & Kovács, G. (2021). Evaluating the evidence for expectation suppression in the visual system. *Neuroscience & Biobehavioral Reviews, 126*, 368–381. https://doi.org/10.1016/j.neubiorev.2021.04.002

Fiser, J., & Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science, 12*(6), 499–504. https://doi.org/10.1111/1467-9280.00392

Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: The paradox of statistical learning. *Trends in Cognitive Sciences, 19*(3), 117–125. https://doi.org/10.1016/j.tics.2014.12.010

Han, B., Mostert, P., & de Lange, F. P. (2019). Predictable tones elicit stimulus-specific suppression of evoked activity in auditory cortex. *NeuroImage, 200*, 242–249. https://doi.org/10.1016/j.neuroimage.2019.06.033

Hindy, N. C., Ng, F. Y., & Turk-Browne, N. B. (2016). Linking pattern completion in the hippocampus to predictive coding in visual cortex. *Nature Neuroscience, 19*(5), Article 5. https://doi.org/10.1038/nn.4284

Jolly, E. (2018). Pymer4: Connecting R and Python for linear mixed modeling. *Journal of Open Source Software, 3*(31), 862. https://doi.org/10.21105/joss.00862

van Kesteren, M. T. R., Ruiter, D. J., Fernández, G., & Henson, R. N. (2012). How schema and novelty augment memory formation. *Trends in Neurosciences, 35*(4), 211–219. https://doi.org/10.1016/j.tins.2012.02.001

Kim, R., Seitz, A., Feenstra, H., & Shams, L. (2009). Testing assumptions of statistical learning: Is it long-term and implicit? *Neuroscience Letters, 461*(2), 145–149. https://doi.org/10.1016/j.neulet.2009.06.030

Kok, P., Jehee, J. F. M., & de Lange, F. P. (2012). Less is more: Expectation sharpens representations in the primary visual cortex. *Neuron, 75*(2), 265–270. https://doi.org/10.1016/j.neuron.2012.04.034

Kok, P., Rait, L. I., & Turk-Browne, N. B. (2019). Content-based dissociation of hippocampal involvement in prediction. *Journal of Cognitive Neuroscience, 32*(3), 527–545. https://doi.org/10.1162/jocn_a_01509

Kok, P., & Turk-Browne, N. B. (2018). Associative prediction of visual shape in the hippocampus. *Journal of Neuroscience, 38*(31), 6888–6899. https://doi.org/10.1523/JNEUROSCI.0163-18.2018

Kumar, S., Kaposvari, P., & Vogels, R. (2017). Encoding of predictable and unpredictable stimuli by inferior temporal cortical neurons. *Journal of Cognitive Neuroscience, 29*(8), 1445–1454. https://doi.org/10.1162/jocn_a_01135

de Lange, F. P., Heilbron, M., & Kok, P. (2018). How do expectations shape perception? *Trends in Cognitive Sciences, 22*(9), 764–779. https://doi.org/10.1016/j.tics.2018.06.002

Manahova, M. E., Mostert, P., Kok, P., Schoffelen, J.-M., & De Lange, F. P. (2018). Stimulus familiarity and expectation jointly modulate neural activity in the visual ventral stream. *Journal of Cognitive Neuroscience, 30*(9), 1366–1377. https://doi.org/10.1162/jocn_a_01281

Meyer, T., & Olson, C. R. (2011). Statistical learning of visual transitions in monkey inferotemporal cortex. *Proceedings of the National Academy of Sciences, 108*(48), 19401–19406. https://doi.org/10.1073/pnas.1112895108

Mitchel, A. D., & Weiss, D. J. (2011). Learning across senses: Cross-modal effects in multisensory statistical learning. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 37*(5), 1081–1091. https://doi.org/10.1037/a0023700

Oliva, A., & Torralba, A. (2007). The role of context in object recognition. *Trends in Cognitive Sciences, 11*(12), 520–527. https://doi.org/10.1016/j.tics.2007.09.009

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., … Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods, 51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Pinto, Y., van Gaal, S., de Lange, F. P., Lamme, V. A. F., & Seth, A. K. (2015). Expectations accelerate entry of visual stimuli into awareness. *Journal of Vision, 15*(8), 13. https://doi.org/10.1167/15.8.13

Richter, D., & de Lange, F. P. (2019). Statistical learning attenuates visual activity only for attended stimuli. *eLife, 8*, 1–27. https://doi.org/10.7554/eLife.47869

Richter, D., Ekman, M., & de Lange, F. P. (2018). Suppressed sensory response to predictable object stimuli throughout the ventral visual vtream. *The Journal of Neuroscience, 38*(34), 7452–7461. https://doi.org/10.1523/JNEUROSCI.3421-17.2018

Richter, D., Heilbron, M., & de Lange, F. P. (2022). Dampened sensory representations for expected input across the ventral visual stream. *Oxford Open Neuroscience, 1*. https://doi.org/10.1093/oons/kvac013. kvac013.

Rosenthal, C. R., Mallik, I., Caballero-Gaudes, C., Sereno, M. I., & Soto, D. (2018). Learning of goal-relevant and -irrelevant complex visual sequences in human V1. *NeuroImage, 179*, 215–224. https://doi.org/10.1016/j.neuroimage.2018.06.023

Saffran, J. R. (2002). Constraints on statistical language learning. *Journal of Memory and Language, 47*(1), 172–196. https://doi.org/10.1006/jmla.2001.2839

Seitz, A. R., Kim, R., van Wassenhove, V., & Shams, L. (2007). Simultaneous and independent acquisition of multisensory and unisensory associations. *Perception, 36*(10), 1445–1453. https://doi.org/10.1068/p5843

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica, 30*, 276–315. https://doi.org/10.1016/0001-6918(69)90055-9

Turk-Browne, N. B., Jungé, J., & Scholl, B. J. (2008). Multidimensional visual statistical learning. *Journal of Experimental Psychology. General, 134*(4), 552–564. https://doi.org/10.1037/0096-3445.134.4.552

Turk-Browne, N. B., Jungé, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General, 134*(4), 552–564. https://doi.org/10.1037/0096-3445.134.4.552

Turk-Browne, N. B., Scholl, B. J., Johnson, M. K., & Chun, M. M. (2010). Implicit perceptual anticipation triggered by statistical learning. *The Journal of Neuroscience, 30*(33), 11177–11187. https://doi.org/10.1523/JNEUROSCI.0858-10.2010

Wyart, V., Nobre, A. C., & Summerfield, C. (2012). Dissociable prior influences of signal probability and relevance on visual contrast sensitivity. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.1120118109

Yon, D., Gilbert, S. J., de Lange, F. P., & Press, C. (2018). Action sharpens sensory representations of expected outcomes. *Nature Communications, 9*(1), 4288. https://doi.org/10.1038/s41467-018-06752-7