The effects of expectation and task-relevance on sensory processing in conscious perception

Lagerman, Jet

Citation

Lagerman, J. (2021). *The effects of expectation and task-relevance on sensory processing in conscious perception*.

The effects of expectation and task-relevance on sensory processing in conscious perception

Jet (J.) Lageman
s1653849
jlageman@live.nl
Institute of Psychology, Leiden University
31-7-2021

External supervisors: Simon van Gaal & Johannes Fahrenfort
First corrector: Sander Nieuwenhuis
Second corrector: Sebo Uithol

Abstract

Identical sensory input may be perceived differently, based on expectations and goals. For example, object recognition is facilitated for expected or task-relevant objects. At the same time, unexpected objects are found to elicit a stronger neural response. These effects can be explained by predictive coding accounts of visual processing, presenting perception as a process of minimizing the difference between predicted and observed sensory input. However, as expectation and task-relevance are often conflated, it is insufficiently clear how these factors influence sensory processing in conscious perception. The current study aimed to investigate the relative influence of expectation and task-relevance on behavioral and neural measures of perception. During two EEG sessions, participants performed a task in which they discriminated between masked face and house images while we independently manipulated expectation and task-relevance. We find that images were more often correctly recognized when they were expected or task-relevant. In addition, we used multivariate pattern analysis to show that a classifier trained on sensory representations of face and house stimuli is better able to distinguish between expected or task-relevant face and house images than between unexpected or task-irrelevant images. These results suggest that expectation and task-relevance have independent effects on sensory processing. Finally, our results show that cue-based manipulations may activate sensory templates even before stimulus onset. We therefore recommend that future studies manipulate expectation and task-relevance without the use of explicit cues.

Keywords: visual perception; expectation; task-relevance; electroencephalography; predictive coding; consciousness

Introduction

Looking at a cluttered desk, you may see your laptop, a pen that you always carry around, or the set of keys you are looking for. This is an example of how identical sensory input may be perceived differently based on our expectations and goals. We are influenced by statistical regularities in the outside world, converted by our brains into predictions or expectations that facilitate object processing and recognition for objects in a familiar context (Biederman, Mezzanotte, & Rabinowitz, 1982; Davenport & Potter, 2004). For example, it is easier to recognize a sink when it is surrounded by kitchen objects than when it is surrounded by garden objects, because we expect sinks to be present in kitchens. At the same time, unexpected or surprising objects may elicit a stronger neural response (Meyer & Olson, 2011), while neural responses to expected objects may be attenuated due to increased encoding efficiency (Kok, Jehee, & de Lange, 2012a; Summerfield & de Lange, 2014; Summerfield et al., 2008).

These seemingly contradictive effects of expectation may be explained by predictive coding accounts of visual processing. According to predictive coding theories, the brain can be seen as a hypothesis-testing machine which is constantly updating a hierarchical model of the world by minimizing the precision-weighted difference between predicted and observed sensory input (Friston, 2009, 2010; Hohwy, 2012). This process would consist of recurrent hierarchical message passing between so-called *representational units* whose activity encodes predictions from the model and *error units* whose activity encodes prediction errors or the difference between predicted and observed sensations (Friston, 2003, 2005, 2008). Fitting with the hierarchical organization of the visual cortex, these units are proposed to exchange information hierarchically. The representational units send messages to units on the same level or lower levels in the hierarchy through backward connections, while error units send messages to units on the same level or higher levels in the hierarchy through forward connections (Friston, 2008). As a result of this communication process, the internal model of the world is improved to resolve the discrepancies between predictions and input.

Predictive coding theories, often presented as ‘unified’ brain theories, offer an explanation of information processing in the brain that supports theories of perception, action, and learning (Clark, 2013). However, critics note that this view does not fully account for the impact of salience or motivational relevance on perception (Bowman et al., 2013). Recall the previous example of the sink, but now imagine that you are thirsty and looking for a source of water. In this case you would probably recognize the sink quite easily, even when it is surrounded by garden objects, because it is motivationally relevant. During any task, the brain may selectively attend to objects that inform action selection, increasing the precision of task-relevant sensory data and thereby facilitating perception of the object (Hohwy, 2012; Parr & Friston, 2019).

Focusing on behavioral measures of perception, it is often very difficult to tell whether an effect has been caused by prior expectations or by task-relevance. Did you notice your keys on the cluttered desk because you had expected that you left them there, or because they are relevant to the task ‘finding your keys’? Another complicating factor is that studies on expectation often use explicit cues which are similar to cues used to direct selective attention to task-relevant stimuli (Rungratsameetaweemana et al., 2018). This causes confusion even in studies using neural measures such as functional MRI (fMRI) or electroencephalography (EEG). Accordingly, results on the effects of expectation and task-relevance on perception are mixed, with studies finding interacting effects (Kok et al., 2012b; Jiang, Summerfield, & Egner, 2013; Smout et al., 2019; St. John-Saaltink et al., 2015) or no interacting effects (Kok et al., 2012a; Sherwell, Garrido, & Cunnington, 2017; Wyart, Nobre, & Summerfield, 2012), and early modulation (Kok et al., 2017; Wyart et al., 2012) or late modulation (Alilović et al., 2019; Rungratsameetaweemana et al., 2018) of perceptual processing. Thus, it is insufficiently clear how (and if) expectation and task-relevance interact during visual processing. This calls for an experiment in which both factors are independently manipulated, so that we can study their relative impact on perception.

Another question that is still to be answered is at which point in time perceptual processing is influenced by expectation and task-relevance, respectively. A common hypothesis, informed by predictive coding theory, is that expectation modulates top-down signaling while task-relevance modulates bottom-up signaling (Gordon et al., 2019). Friston (2005) notes that forward connections in the brain are mediated by faster AMPA and $GABA_A$ receptors, while backward connections are mostly mediated by slower NMDA receptors. In the case that expectation and task-relevance have differential effects on the activity of representational units and error units, then we may also expect the impact of these factors to vary over time. However, a large portion of previous research has been conducted with fMRI (e.g., Summerfield et al., 2008; Egner, Monti, & Summerfield, 2010; Kok et al., 2011; Kok et al., 2012; Jiang et al., 2013), which does not have sufficient temporal precision to detect activity changes on the millisecond-level. EEG does have the temporal precision to detect these millisecond-level changes, but so far, even studies conducted with EEG are not fully conclusive. This could be explained by differences between studies in task design or analytic approach (Slagter, Alilović, & van Gaal, 2017). In the current study, we aim to examine the temporal dynamics of sensory processing by analyzing EEG data with multivariate pattern analysis.

Finally, sensory processing of a stimulus is not equivalent to consciously accessing that stimulus. Not all input that is processed may reach consciousness. Theoretical accounts propose that conscious perception consists of the brain's 'best guess' about the causal structure of the world, or the internal model that best succeeds in minimizing prediction error (Hohwy, 2012). However, other models that do not enter consciousness may still be able to elicit and suppress prediction errors. As we cannot tell the contents of consciousness from brain activity, studies rely on subjective reports to measure whether a presented stimulus was consciously perceived. Using this approach, studies have found that expected stimuli are more likely to be consciously accessed than unexpected stimuli (Pinto et al., 2015; Meijs et al., 2018; Meijs et al., 2019), but findings regarding task-relevance are inconclusive. Adding to subjective reports and EEG data, we included the Perceptual Awareness Scale (PAS; Ramsøy & Overgaard, 2004) in the current study to measure participants' subjective awareness of stimuli. In this way, we aim to explore the link between expectation, task-relevance, and conscious perception.

In sum, the current experiment was designed to study the impact of task-relevance and expectation on sensory processing in conscious perception. In order to do so, participants performed a task in which expectation and task-relevance of face and house stimuli were independently manipulated, while brain activity was measured with EEG. To explore the neural underpinnings of behavior on a millisecond-level, we trained a linear discriminant classifier on an independent training task to discriminate between face and house trials. This allowed us to compare the relative strength of face and house representations as a function of expectation and task-relevance, and for different levels of subjective awareness (PAS scores). Finally, we repeated our analyses for ambiguous stimuli that contained both face and house images, as we attempted to account for initial differences between face and house stimuli. For example, expectation may affect the processing of faces differently than the processing of houses. The use of ambiguous stimuli allowed us to study the effects of expectation and task-relevance independently of the object of perception, that is, the image that is shown.

Methods

Participants

27 healthy, right-handed participants (10 male, 17 female, mean age = 21.67) with normal or corrected-to-normal vision gave informed written consent to participate in this two-session study. Participants were recruited at the University of Amsterdam and participation was rewarded with either 60 euros or course credits. One participant was excluded from all analyses due to a technical failure in the second session.

Tasks

Each session, the participants performed two tasks. During the first, an N-back task, they were asked to watch a stream of continuously presented face and house images, separated by a temporal gap. In most trials, the stimulus was different from the previously presented stimulus. Participants were instructed to press a key whenever a stimulus was repeated (i.e., the exact same picture was shown two consecutive times). This task was included in the design as an independent localizer task, for the purpose of collecting independent training data in order to train a neural classifier to discriminate between ‘face’ trials and ‘house’ trials based on the stimulus-locked preprocessed EEG data. The N-back task is a relatively simple task to ensure that participants pay attention to the face and house images, without any cognitive or perceptual manipulations.

During the second task, which we will refer to as the Orientation Discrimination task, we manipulated expectation and task-relevance, and we measured our variables of interest. The EEG data collected during this task was used to test the performance of the neural classifier trained to discriminate between faces and houses. In this task, participants were first presented with a cue to indicate a task-relevant category, and subsequently with a stimulus of which they were asked to judge the orientation, but *only* if the stimulus category matched the cue (e.g., ‘face’ cue and ‘face’ stimulus). Between blocks, we varied the relative frequency of each stimulus category as a manipulation of participants’ expectation. Since participants were asked to judge the stimulus *orientation* rather than its category, the expectation manipulation should not affect task-relevance. Vice versa, the task-relevance cue indicated which category should be responded to but did not *predict* the category of the following stimulus, and therefore it should not affect participants’ expectations about the upcoming stimulus. In sum, manipulations of task-relevance and expectation were fully independent.

Stimuli

The experimental tasks were programmed and executed using Psychophysics Toolbox (version 3.0.14; Brainard, 1997) within the Matlab environment (R2010, MathWorks). Stimuli were presented on an Asus VG236H LCD monitor (23” diagonal, 1920 x 1080 pixel resolution; 100 Hz refresh rate) at a viewing distance of 60 cm. All stimuli were presented on a grey background (RGB: [128 128 128]).

In the N-back task, target stimuli consisted of 180 house images and 180 face images (90 male, 90 female). Other than during the consecutive repeat trials, no stimulus was shown more than once per block of 180 trials. Both houses and faces were clearly visible. All images were greyscale, sized 16 x 20˚, and centrally presented. For this task, the images were not equalized on luminance or spatial frequency characteristics.

In the Orientation Discrimination task, target stimuli consisted of 180 house images, 180 face images (90 male, 90 female), and 144 ambiguous face-house images. House and face images were shown 1-2 times per session (1.6 times: all 180 stimuli were presented once, then a randomly selected subset of 108 were presented again). The ambiguous images were created by randomly matching each of the face images to a house image, making each stimulus 50% transparent and superimposing them. The process of matching each face image to a house image was repeated five times to obtain five sets of 180 ambiguous images. Per participant, one set was selected (counterbalanced across participants) from which 144 images were randomly selected for presentation. Again, all images were greyscale, sized 16 x 20˚, and centrally presented at an orientation of 10˚ angular deviation either clockwise (‘right’) or anticlockwise from vertical (‘left’). They were immediately preceded and followed by scrambled masks presented for 50 ms. These masks were sized 22 x 22˚ and were randomly selected from 900 images of scrambled face and house images. Masks were created by selecting face and house images and parsing these into 12 x 15 tiles. Tiles were then randomly shuffled, made transparent and superimposed to create a scrambled image. The SHINE toolbox (Williams, Willenbockel, & Gauthier, 2009) was used to equate all face, house and ambiguous stimuli for spatial frequency (equating rotational average of the Fourier amplitude spectrum) and then for luminance (equating luminance histograms) over 20 iterations.

For both tasks, all house stimuli were obtained from Egner, Monti, and Summerfield (2010), and face stimuli were compiled from the Cohn-Kanade Facial Expression Database (Kanade, Cohn, & Tian, 2000) and from Endl and colleagues (1998).

Experimental procedure

The study consisted of two identical EEG sessions in which the participants first performed an N-back task and subsequently performed an Orientation Discrimination task. In the N-back task (Figure 1A), participants started with 40 practice presentations, then 540 stimulus presentations per experimental session. These were divided over 3 blocks of 180 stimulus presentations. Target stimuli (face or house images) were continuously presented for 200 ms, separated by a gap of 1200-1400 ms. Every 8-12 presentations (randomized), a stimulus image was repeated. That is, the same image was shown two consecutive times. Participants were instructed to press the space bar to indicate that the current stimulus was a repeat of the previous stimulus. Accuracy feedback was provided to the participants at the end of each block, with specific additional feedback if they had made more than 2 mistakes (*miss:* not pressing the spacebar when a stimulus was repeated; *false alarm:* pressing the spacebar when a stimulus was not repeated).

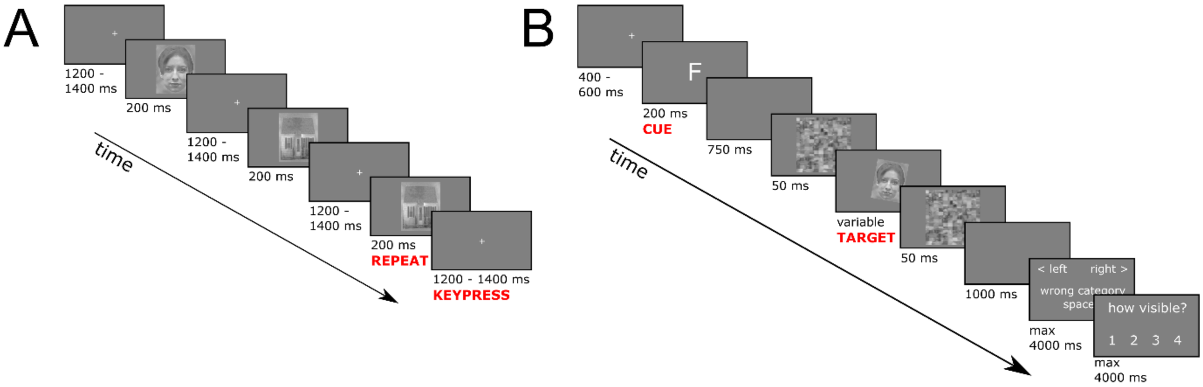


Figure 1. Timeline of the experimental tasks. **A)** shows two trials of the N-back task, used for training a linear discriminant classifier to distinguish face images from house images. **B)** shows one trial of the Orientation Discrimination task on which we tested the classifier.

In the Orientation Discrimination task (Figure 1B), the two key manipulations were *task relevance* (task-relevant vs. task-irrelevant) and *expectation* (expected vs. unexpected). Task relevance was manipulated trial-wise by cuing a stimulus category for which to make a tilt discrimination, using either the letter F (for ‘face’) or H (for ‘house’). Expectation was manipulated block-wise by varying the relative frequencies of each stimulus category as the target image (60% expected category, 20% unexpected category, 20% ambiguous). In each session, participants completed 720 trials, divided into 12 blocks of 60 trials each. Before starting each block, participants were explicitly informed as to whether they could expect to see more houses or faces in that block. The expected category alternated between blocks, with odd-numbered participants starting with ‘face’ and even-numbered participants starting with ‘house’ as the expected category. All other experimental conditions, such as cue letter and tilt orientation, were counterbalanced within blocks, making all options equally likely across trials with expected, unexpected, and ambiguous stimuli.

Each trial began with a fixation point which lasted for 400-600 ms. Subsequently, a cue letter which indicated the relevant stimulus category (‘F’ for face / ‘H’ for house; in white 128-point Arial font, 2.7 x 1.8°) was presented for 200 ms. Then, 800 ms after the cue, the target stimulus was presented for a variable duration (see section below for duration titration procedure), immediately preceded and followed by a scrambled mask that was presented for 50 ms. A response screen was presented 1000 ms after the second mask offset, instructing participants to indicate whether the presented image was tilted to the left or right if from the cued stimulus category, or to indicate ‘wrong category’ if the presented image was not of the cued category (‘spacebar’ keyboard response). By including the ‘wrong category’ option, the participant’s answer contained information on both the perceived stimulus category and the perceived tilt direction of task-relevant stimuli. That is, if participants made a tilt judgment, they implicitly indicated that they believed the stimulus to be of the same category of the cue. In contrast, if they pressed the spacebar, they indicated that they believed the stimulus to be of the category opposite to the cue. For example, after a ‘face’ cue, a ‘left’ answer could be interpreted as a ‘face’ category judgment, while a spacebar press could be interpreted as a ‘house’ category judgment. To prevent motor response preparation prior to the response screen, stimulus-response mappings were randomly determined on each trial and indicated by presenting the words ‘left’ and ‘right’ either on the left (indicating left-handed keyboard response: ‘z’/‘x’) or right side (indicating right-handed keyboard response: ‘,<’/‘.>’) of the response screen. Finally, participants viewed a second response screen which instructed them to report on their experience of the target image using the PAS visibility scale (Ramsøy & Overgaard, 2004; 1: no experience, 2: brief glimpse, 3: almost clear experience, 4: clear experience). Trials ended if no response was recorded within 4000 ms of either response screen.

The exact duration of the target stimulus during this task was determined by the current participant’s performance on the category judgments, quantified by the sensitivity score d’ (see section ‘Behavioral analysis’ for calculation details). During practice trials and during the first experimental block, all stimuli were presented for 30 ms. After the first block, stimulus duration was decreased by 10 ms if a participant obtained a score of d’ > 1.5 in the previous block, until a lower limit of 10 ms. If the participant obtained a score of d’ < 1.0 in the previous block, stimulus duration was increased by 10 ms. Within blocks, stimulus duration remained constant.

Behavioral analysis

Performance scores for the Orientation Discrimination task were calculated on the basis of the category judgment response to face and house target images. Responses were scored as hits (correct category judgment, e.g., responding 'face' to a face stimulus) or false alarms (incorrect category judgment, e.g., responding 'house' to a face stimulus).

The hit rate (HR) and false alarm rate (FAR) were then calculated as the proportion of hits and the proportion of false alarms, respectively for each of the response options. For example, the HR for the 'house' response would be computed as the number of times a participant responded 'house' to a *house* stimulus, divided by the total number of 'house' responses. FAR for the 'house' response would be computed as the number of times a participant responded 'house' to a *face* stimulus, divided by the total number of 'house' responses. Sensitivity (d') was finally computed as HR minus FAR for each of the four combinations of expectation and task-relevance: expected and task-relevant, expected and task-irrelevant, unexpected and task-relevant, or unexpected and task-irrelevant. We performed a repeated-measures ANOVA to test the main and interaction effects of expectation and task-relevance on d'.

Note that by comparing these conditions, trials are considered outside of their natural context, as an 'expected' house stimulus occurs only during blocks where houses are expected while an 'expected' face stimulus occurs only during blocks where faces are expected. Thus, we selected trials from blocks where faces were expected and blocks where houses were expected, and computed performance on these trial selections. This allows us to selectively quantify how well participants can differentiate between expected faces and expected houses, or how well they can differentiate between unexpected faces and unexpected houses, and similarly for task-relevant faces or houses, and task-irrelevant faces or houses. Naturally, performance shifts can in this case be explained by bias. We explicitly investigate those bias effects in a set of independent analyses on the ambiguous stimuli.

For the ambiguous stimuli, we calculated the proportion of 'face' category judgments for each of the four combinations of expected category (face vs. house) and cue type (face vs. house). Since each ambiguous target image contained both a face image and a house image, stimuli would always be both expected and unexpected, and both task-relevant and task-irrelevant. Likewise, responses could not be classified as correct or incorrect. However, we could investigate whether participants were more or less biased towards a ‘face’ or ‘house’ response under each of the manipulations.

Finally, we looked at category judgments and (where possible) sensitivity for each level of subjective visibility as measured by the PAS visibility scale (Ramsøy & Overgaard, 2004; 1: no experience, 2: brief glimpse, 3: almost clear experience, 4: clear experience).

EEG data collection and preprocessing

EEG data were collected at 512 Hz using a 64-channel ActiveTwo system (BioSemi). Data preprocessing was performed using Matlab (R2019b, MathWorks), the EEGLAB toolbox (Delorme & Makeig, 2004), and the ADAM toolbox (Fahrenfort, van Driel, van Gaal, & Olivers, 2018). EEG data collected during the N-back task and Discrimination Orientation task were saved in separate data files and preprocessed separately.

For the N-back task, all data were first referenced to the average of the earlobes and down-sampled to 128 Hz. Subsequently, we applied trial-masked robust detrending with a polynomial order of 30 and masking period of 0 to 1000 ms relative to stimulus onset, using the detrend_and_epoch function in the ADAM toolbox (van Driel, Olivers, & Fahrenfort, 2021). The continuous dataset was epoched between -200 and 1000 ms relative to stimulus onset, before running independent component analysis (ICA) to detect and remove eyeblink artifacts. The data from both experimental sessions were then merged into one file per participant. Finally, we applied baseline correction using a baseline from -200 to 0 ms relative to stimulus onset.

For the Orientation Discrimination task, all data were first referenced to the average of the earlobes and down-sampled to 128 Hz. We applied trial-masked robust detrending with a polynomial order of 30 and masking period of -1000 to 1000 ms relative to stimulus onset. This masking period includes the time between cue onset and stimulus onset, as we were also interested in the EEG signal during this period. The continuous dataset was then epoched between -1400 and 2000 ms relative to stimulus onset. We chose the length of this epoch specifically to include the first fixation screen and the task-relevance cue in each epoch. Again, we ran ICA to detect and remove eyeblink artifacts and merged the data from both experimental sessions into one file per participant. We also applied baseline correction using a baseline from -1200 to -1000 ms relative to stimulus onset (-200 to 0 ms relative to cue onset).

EEG multivariate pattern analysis

For each participant, we applied the backward decoding classification algorithm from the ADAM toolbox (Fahrenfort et al., 2018), using the data from the N-back task for training and the data from the Orientation Discrimination task for testing. We trained a linear discriminant classifier to discriminate face and house target stimuli based on the raw EEG activity across all 64 electrodes in the training data. To separate sensory processes from decision processes, we used an approach similar to Mostert, Kok, & de Lange (2015). Instead of computing the full time generalization matrix, we opted to train the classifier only on the peak decoding time ± 5 ms. This timepoint was determined by first training and testing the classifier only on the N-back task data using 5-fold cross-validation, and subsequently determining the timepoint for which decoding accuracy was maximized. This resulted in a training window of 128-138 ms relative to stimulus onset. Our reasoning behind the use of this training window was twofold. First, the advantage of an independent, not decision-related training set is that the classifier cannot be guided by task-specific or decision-related patterns in the testing set. To maximize the benefits following from this advantage, we should consider off-diagonal decoding (i.e., in which decoding is trained on one timepoint and tested on another timepoint), which allows us to assume that accurate decoding reflects neural patterns shared between the training and testing set (King & Dehaene, 2014). In this way, we hope to ensure that accurate decoding reflects sensory processes and not decision-related processes or pure coincidence. Second, more practically, training on a smaller time-window instead of computing the full time generalization matrix greatly limits computation time.

Next, we tested the classifier and computed classification accuracy as the area under the curve (AUC; Bradley, 1997) for each participant and for each combination of expectation and task-relevance. As in the behavioral analyses, this means that trials were not considered in their natural context, but compared to trials from different blocks. Note that all analyses involving AUC exclude ambiguous target stimuli from the testing data, as these do not have a veridical label to classify. We will later describe our separate analysis of the ambiguous stimuli. On all decoding output, two-sided significance testing was performed using cluster based permutation testing with 1000 permutations. However, note that this testing informs us only of whether decoding accuracy is significantly different from chance-level, and not of whether it is different from decoding in other conditions.

In order to compare decoding between conditions, we extracted the AUC values for each condition at 172 ms after stimulus onset. This timepoint was determined by averaging decoding across all four conditions and choosing the timepoint at which decoding accuracy was maximized. Because we determined the peak of the average decoding across conditions, its temporal location is independent of condition. Therefore, it should also be independent of our manipulations. For the extracted AUC values, we performed repeated-measures ANOVA to test the main and interaction effect of task-relevance and expectation on decoding accuracy. Because we also wanted to visualize the main effects over time, we then averaged the decoding from the two task-relevant conditions (expected, task-relevant and unexpected, task-relevant) and the two task-irrelevant conditions (expected, task-irrelevant and unexpected, task-irrelevant), and computed their difference score. The averaging is meant to obtain the individual effect of task-relevance, regardless of expectation. We did the same for the two expected conditions (expected, task-relevant and expected, task-irrelevant) and the two unexpected conditions (unexpected, task-relevant and unexpected, task-irrelevant), and also computed their difference score. The difference scores should be a visualization of the main effects of task-relevance and expectation over time.

Then, we analyzed the results separately for each level of the PAS visibility scale (1: no experience, 2: brief glimpse, 3: almost clear experience, 4: clear experience). In order to achieve this, the trained linear discriminant classifier was tested separately for each participant, each condition, and each PAS level. Because not every participant answered all PAS scores in each condition, the results of this analysis procedure may be based on a lower number of participants. This is clearly indicated in each figure in the Results section. We first plotted the average decoding across all conditions for each PAS level. In order to investigate the relationship between subjective awareness and decoding, we extracted the AUC values at 172 ms after stimulus onset to perform repeated-measures ANOVA.

Finally, we analyzed the testing trials with ambiguous stimuli. For these trials, it is not possible to compute classification accuracy or AUC, since there is not a ‘true’ label. However, it is possible to see whether a classifier trained to discriminate faces and houses is more likely to give a ‘face’ or ‘house’ label at any timepoint during the ambiguous trial. Measuring the probability of a ‘face’ label instead of the AUC, we performed the same analyses as for the face/house trials. First, we computed the probability of a ‘face’ label for each of the combinations of expectation and task-relevance and performed repeated-measures ANOVA on the value at 172 ms after stimulus onset. Second, we visualized the main effects of task-relevance and expectation over time by averaging and computing difference scores. Third, we plotted the average probability across all conditions for each PAS level, again performing repeated-measures ANOVA on the value at 172 ms after stimulus onset.

Results

Figure 2 shows the behavioral responses to face and house target stimuli in the Orientation Discrimination task for each combination of expectation and task-relevance. In Figure 2A, we see that quite reasonably, face stimuli were more often recognized as faces than house stimuli, while house stimuli were more often recognized as houses than face stimuli. Moreover, for both face and house stimuli, participants were more likely to give a ‘face’ response when they expected to see more faces in that block (face stimuli: $F(1, 25) = 41.26$, $p < 0.001$; house stimuli: $F(1, 25) = 17.75$, $p < 0.001$). Likewise, for both face and house stimuli, participants also responded ‘face’ more often when faces were the task-relevant category in that trial (face stimuli: $F(1, 25) = 9.72$, $p = 0.005$; house stimuli: $F(1, 25) = 47.33$, $p < 0.001$). We did not find an interacting effect of expectation and task-relevance on the proportion of ‘face’ responses for either stimulus category (face stimuli: $F(1, 25) = 0.00$, $p = 0.961$; house stimuli: $F(1, 25) = 1.39$, $p = 0.249$). Importantly, these results do not tell us whether expectation and task-relevance improve perception of faces and houses, or simply bias participants towards a certain response.

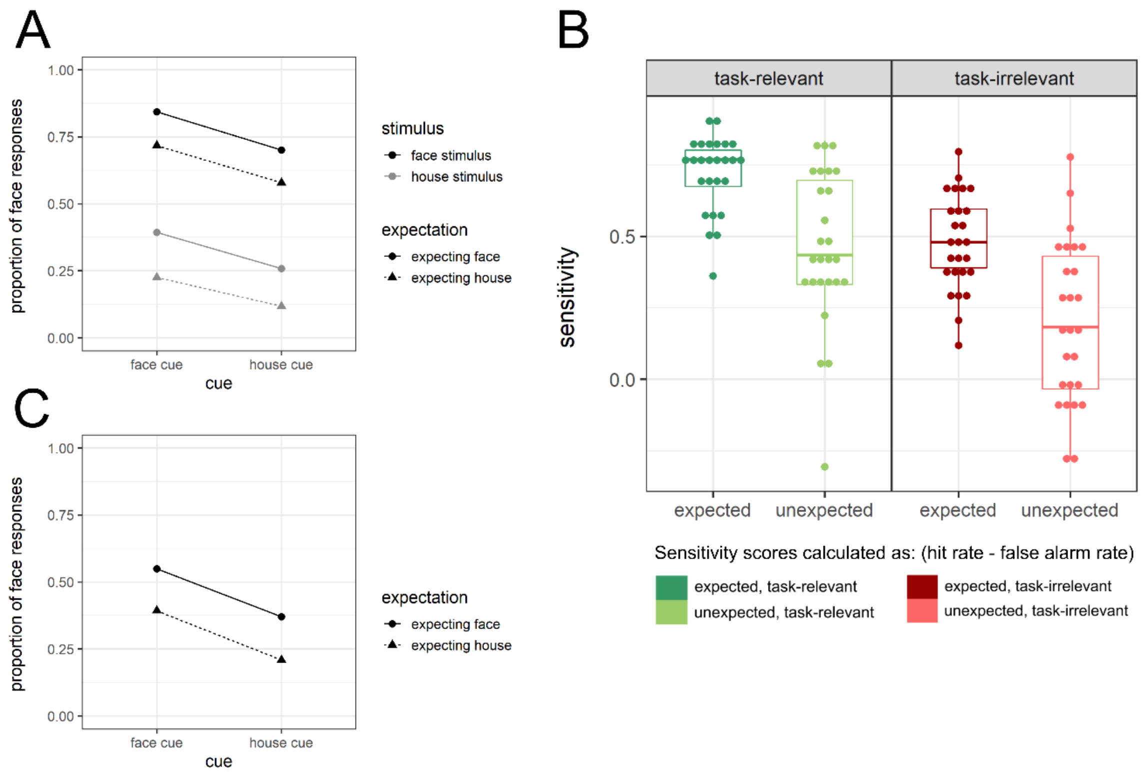


Figure 2. Behavioral results. **A)** shows the average proportion of category responses for face and house stimuli in each condition. On the Y-axis is the proportion of ‘face’ responses. A higher score indicates more ‘face’ responses, while a lower score indicates more ‘house’ responses. **B)** shows the sensitivity scores of each participant (each dot represents one participant) in each condition. See section ‘Behavioral analysis’ in Methods for calculation details of the hit rate and false alarm rate. **C)** shows the average proportion of category responses for ambiguous stimuli in each condition. Again, a higher score indicates more ‘face’ responses, while a lower score indicates more ‘house’ responses.

To show that the effects of expectation and task-relevance are perceptual, it is more informative to inspect the sensitivity scores (Figure 2B). A repeated-measures (RM) ANOVA indicates that both expectation (*F*(1, 25) = 24.38, *p* < 0.001) and task-relevance (*F*(1, 25) = 31.69, *p* < 0.001) have a significant effect on sensitivity, in the absence of an interaction effect (*F*(1, 25) = 0.95, *p* = 0.340). These results show that perceptual sensitivity to both expected stimuli and task-relevant stimuli was higher. However, it is important to recall that these sensitivity scores are not separated by stimulus category and thus force us to make an uncommon comparison. For instance, the 'expected and task-relevant' condition contains both expected face stimuli and expected house stimuli, which can never occur during the same block as expectation is manipulated block-wise. In addition, sensitivity is calculated here as the hit rate in a certain condition minus the false alarm rate. This measure could still be subject

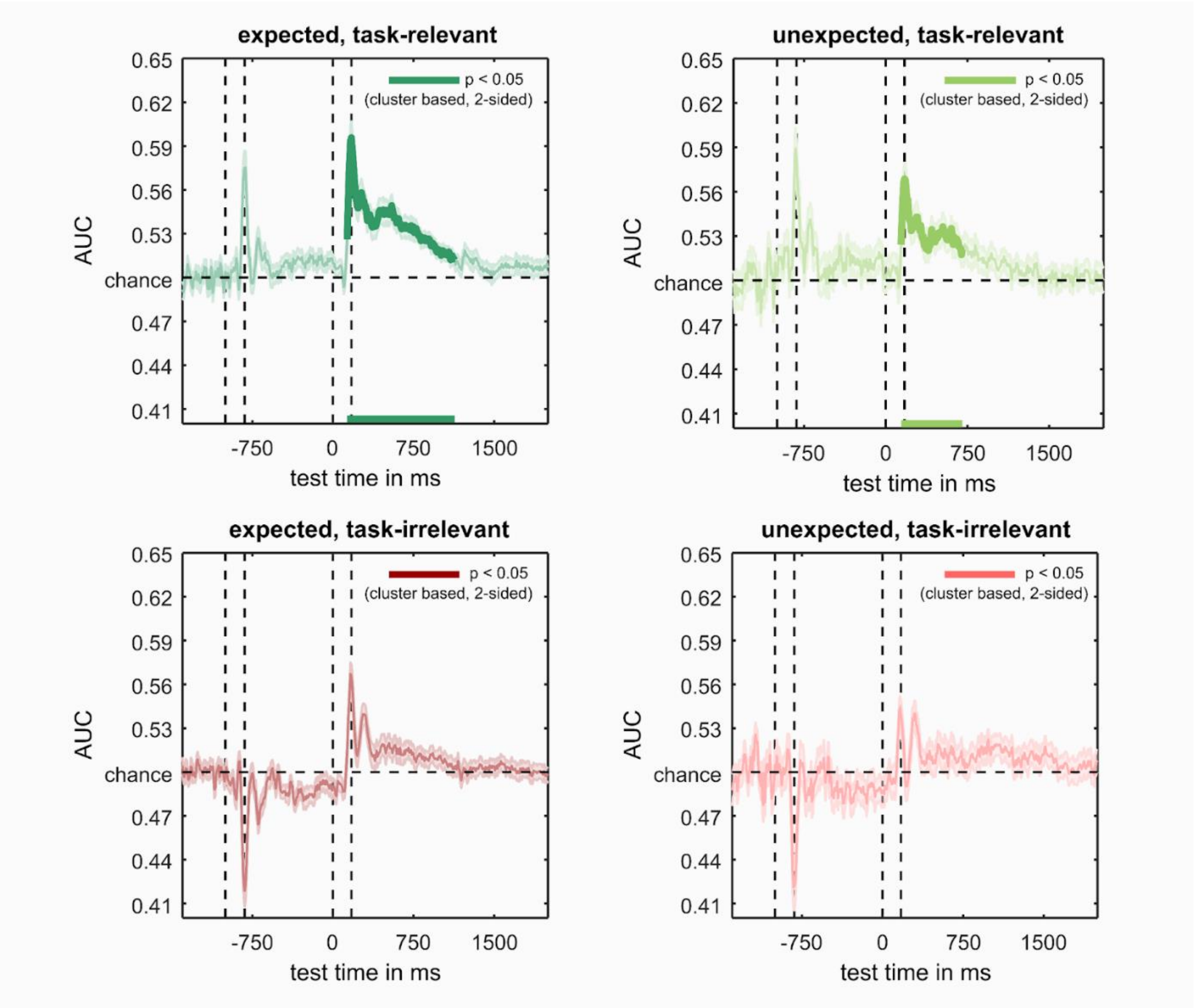


Figure 3. EEG decoding accuracy over time for the four experimental conditions, trained and tested on all electrodes to discriminate face and house target stimuli. Line graphs reflect the average area under the curve (AUC; Bradley, 1997) across participants ± SEM (light colored). Dark bold lines indicate $p < 0.05$ with cluster based permutation. From left to right, dotted vertical lines represent cue onset (-1000 ms), peak cue decoding (-820 ms), target onset (0 ms), and peak target decoding (172 ms).

to (non-perceptual) biases and cannot be considered a bias-free measure of perceptual sensitivity. Therefore, the results in Figure 2B should be interpreted with appropriate caution.

Using only behavioral measures, we are not able to tell how expectation and task-relevance affect sensory processing in the brain. Therefore, we inspected the performance of a linear discriminant classifier trained to decode face and house trials from a participant's EEG data across 64 electrodes. Figure 3 shows how well this classifier performs on average in each of the four conditions. Decoding of the stimulus category is above chance in all conditions, but only significantly so in the 'expected and task-relevant' and 'unexpected and task-relevant' conditions. Decoding peaks on average at 172 ms after stimulus onset, which is represented by the rightmost dotted vertical line. Noteworthy is how classifier performance directly after the task-relevance cue (letter 'F' for face or 'H' for house) quickly rises in task-relevant conditions (upper row; green-colored) and drops in task-irrelevant conditions (bottom row; red-colored). Decoding of the task-relevance cue peaks (either positively or negatively) on average at -820 ms relative to stimulus onset. The opposite direction of the peaks is readily explained by the fact that in task-relevant conditions, the cue is matching the veridical stimulus label, while in task-irrelevant conditions, the cue is the opposite from the veridical stimulus label. Nevertheless, it remains striking that the presentation of a single letter is sufficient to activate the representation of a face or a house, derived from an independent localizer task.

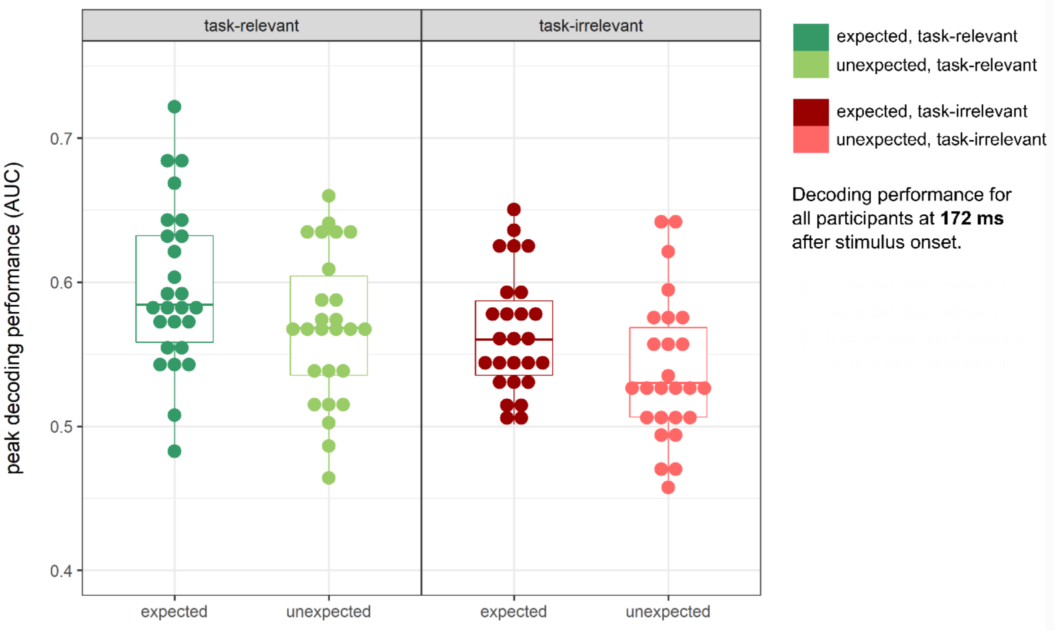


Figure 4. Extracted peak AUC (values at 172 ms after stimulus onset) for each participant in each condition (each dot represents one participant).

In order to statistically compare peak AUC values between conditions, we extracted the decoding performance at 172 ms after stimulus onset (see section 'EEG multivariate pattern analysis' in methods for information on how this timepoint was determined) for each participant. Figure 4 shows these values for each condition. Using RM-ANOVA, we find a main effect of expectation ($F(1, 25) = 7.77$, $p = 0.010$) and task-relevance ($F(1, 25) = 11.31$, $p = 0.002$), but no significant interaction effect ($F(1, 25) = 0.01$, $p = 0.924$). Again, note that we are comparing responses to stimuli which are taken out of their context as to avoid, for example, comparing *expected* faces to *unexpected* houses.

Although it is practical for statistical analysis to compare results extracted from one specific timepoint, for substantive interpretation it may be interesting to consider results over time. Moreover, analyses containing data from only one timepoint are always at risk of being somewhat arbitrary: differences between conditions at an earlier or later timepoint may differ. Further, part of our research aim was to study these changes over time. Therefore, we also inspected at the classifier performance when we averaged the task-relevant (expected) and task-irrelevant (unexpected) conditions and considered the difference between these averages over time (Figure 5). Figure 5A shows that task-relevance mainly affects decoding in the time before stimulus onset. We see a high-amplitude (but unsignificant) peak right after the cue is presented, likely representing decoding of the cue identity, and a lower-amplitude significant cluster of timepoints slightly before stimulus onset. This could perhaps reflect the

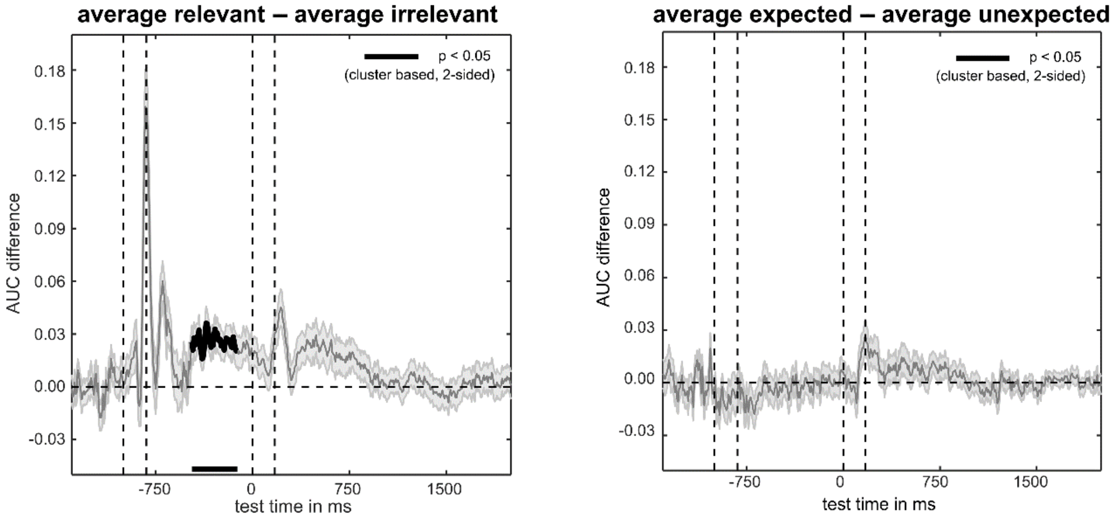
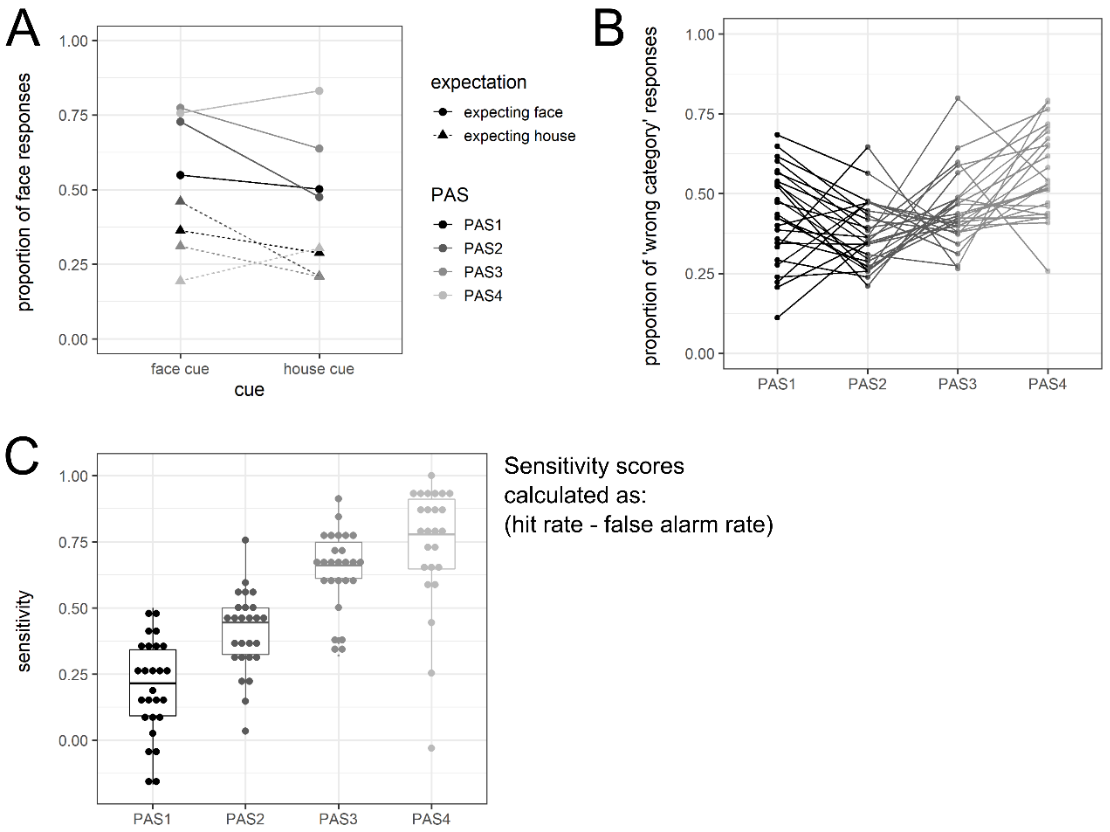


Figure 5. Main effects of task-relevance and expectation on decoding of face and house stimuli. All scores are plotted ± SEM (light colored). Dark bold lines indicate $p < 0.05$ with cluster based permutation. **A)** shows the difference between the average AUC in task-relevant conditions ('expected and task-relevant' and 'unexpected and task-relevant') and the average AUC in task-irrelevant conditions ('expected and task-irrelevant' and 'unexpected and task-irrelevant'). **B)** shows the difference between the average AUC in expected conditions ('expected and task-relevant' and 'expected and task-irrelevant') and the average AUC in unexpected conditions ('unexpected and task-relevant' and 'unexpected and task-irrelevant').

maintenance of task-relevant sensory information (Summerfield & Koechlin, 2008; Weaver et al., 2019). In Figure 5B, we see no such peaks or clusters, implying that there should be no main effect of expectation on the decoding of stimulus labels.

Our next aim was to relate any effects we found of task-relevance and expectation to differences in levels of subjective awareness of stimuli. After all, we do not know whether differences in the neural processing of visual stimuli will translate to differences in subjective reports. In order to investigate to what extent we could link our behavioral and neural measures of perception to subjective awareness, we considered participants' responses to the PAS visibility scale. Given that visibility was not manipulated and that this scale was responded to at the very end of the trial, we should be careful not to assume causality in this set of analyses. As the PAS was responded to only after making a judgment about the tilt and identity of a stimulus, the response could reflect aspects of decision confidence rather than subjective visibility.



Sensitivity scores
calculated as:
(hit rate - false alarm rate)

Figure 6. Behavioral results as a function of the response to the PAS visibility scale. **A)** shows the average proportion of category responses for each PAS level in each condition. On the Y-axis is the proportion of 'face' responses. A higher score indicates more 'face' responses, while a lower score indicates more 'house' responses. **B)** shows the sensitivity scores of 25 participants (each dot represents one participant) for each PAS level. See section 'Behavioral analysis' in Methods for calculation details of the hit rate and false alarm rate. **C)** shows the number of 'wrong category' responses as a proportion of all responses to stimuli ('tilted to the left', 'tilted to the right', or 'wrong category'), as a function of PAS level. Connected dots represent individual participants.

Figure 6 again shows the behavioral results, but separated by PAS level. In Figure 6A we see the proportion of face responses in each condition for all PAS levels. In general, we see the same trends as before: participants give more 'face' responses when they expect to see a face, or when faces are the task-relevant category. Most surprising in this figure is that the relationship between the task-relevance cue type and the number of 'face' responses seems to change in direction for stimuli that participants judged as a 'clear experience' (PAS 4): after a house cue, participants give more 'face' responses than after a face cue. We believe that perhaps, when participants view a target stimulus, the 'default' response is to make a tilt judgment (i.e., indicating that the stimulus category matches the cue), *unless* they are quite certain that the target stimulus does not match the cue category. Logically, this certainty will occur more when participants experience the stimulus more clearly, that is, when they report a higher PAS level. This idea is supported by the results plotted in Figure 6B, where we see that the average proportion of 'wrong category' responses increases with PAS level ($F(3, 22) = 8.43$, $p = 0.001$). Thus, it seems that many participants are more likely to give a 'wrong category' response for stimuli that they experienced clearly (PAS 4). Again, we would like to stress that all results in Figure 6 are correlational, and thus the relationship reversal between task-relevance cue type and response categories may simply reflect this disparity in 'wrong category' responses.

Figure 6C shows that sensitivity increases with increased subjective awareness. This is confirmed by a RM-ANOVA in which we find a clearly significant main effect of PAS level on sensitivity ($F(3, 22) = 31.76$, $p < 0.001$). Our next step was then to investigate whether this increase in sensitivity would also translate to an improvement in decoding accuracy. This would indicate that there is a relationship between the extent to which participants are subjectively aware of a stimulus, and the extent to which we can decode that stimulus from EEG signals. Figure 7 shows the average decoding accuracy (AUC) across all conditions, for each PAS level. At first glance, peak AUC seems to increase with PAS level. However, when we extracted peak AUC values at 172 ms after stimulus onset and performed RM-ANOVA, we did not find a significant effect of PAS level on AUC ($F(3, 12) = 1.32$, $p = 0.313$). It should be noted that this result only included 15 out of the 26 participants, since only these participants had complete results for all PAS levels. Moreover, this test just reflects a single timepoint during the trial and is not indicative of the differences in AUC over time.

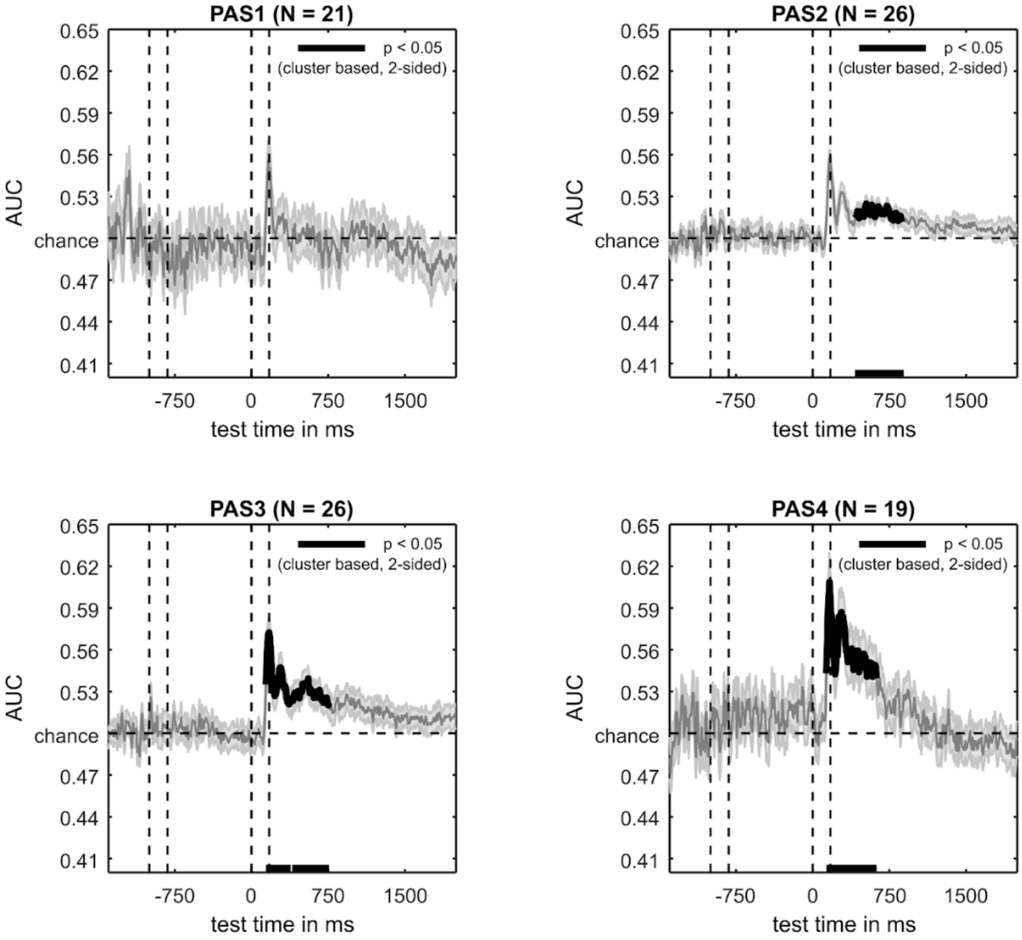


Figure 7. EEG decoding accuracy over time for each PAS level, trained and tested on all electrodes to discriminate face and house target stimuli. Line graphs reflect the average AUC across participants ± SEM (light colored). Dark bold lines indicate $p < 0.05$ with cluster based permutation. From left to right, dotted vertical lines represent cue onset (-1000 ms), peak cue decoding (-820 ms), target onset (0 ms), and peak target decoding (172 ms). Note that the number of participants of which results are averaged is not the same for each subplot. Within each level of the PAS, only participants with complete results (i.e., who responded this PAS level at least once in *each* condition) were considered for this analysis.

In the results so far, we have made an implicit assumption that expectation or task-relevance affect house images in the same way as they affect face images. These results might be affected by prior bias towards 'face' or 'house' responses. Here, the term bias includes any difference between face and house stimuli in perceptual sensation, cognitive processing, or decision-making that is not a consequence of the study manipulations. In order to study the effects of expectation and task-relevance on perception, independently of what is perceived, we repeated our analyses to the ambiguous target stimuli, which contained both a face image and a house image. These analyses also allow us to make a more direct link to consciousness: can a manipulation of expectation or task-relevance change whether participants report seeing a face or a house? To answer this question, we first examine the category responses to ambiguous stimuli, shown in Figure 2C. On average, participants were more likely to classify an ambiguous stimulus as a house than as a face (grand average proportion of 'face' responses: 0.38). We found effects of both expectation (*F*(1, 25) = 53.60, $p < 0.001$) and task-relevance (*F*(1, 25) = 26.75, $p < 0.001$) on response: when faces were task-relevant or expected, the proportion of 'face' responses increased. We did not find an interaction effect (*F*(1, 25) = 0.05, $p = 0.822$). Since there was no 'correct' category response to an ambiguous target stimulus, we were not able to calculate sensitivity for this set of stimuli.

These behavioral results inform us of the possibility that participants may be intrinsically biased towards a 'house' response, but they do not indicate when and how this

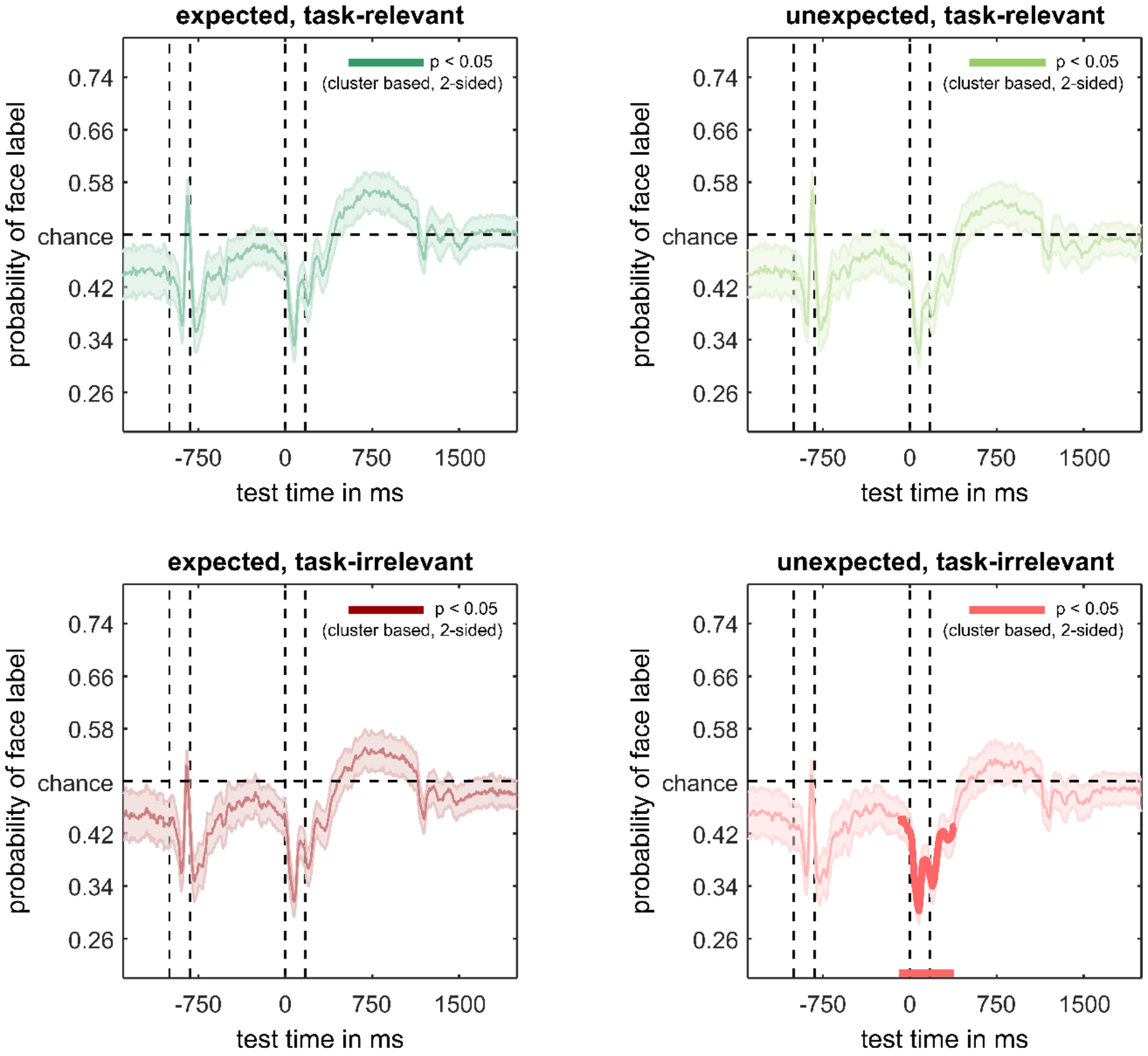


Figure 8. EEG classification over time for the four experimental conditions, trained on all electrodes to discriminate face and house target stimuli, then tested on ambiguous target stimuli. Line graphs reflect the probability of a face label across participants ± SEM (light colored). Higher values indicate that the algorithm is more likely to classify the image as a face, lower values indicate that the algorithm is more likely to classify the image as a house. Dark bold lines reflect $p < 0.05$ with cluster based permutation. From left to right, dotted vertical lines represent cue onset (-1000 ms), peak cue decoding (-820 ms), target onset (0 ms), and peak target decoding (172 ms). Conditions (subplot titles) are phrased from the 'perspective' of faces: 'expected, task-relevant' indicates that *faces* are expected and task-relevant. Likewise, 'unexpected, task-irrelevant' indicates that faces are unexpected and task-irrelevant (i.e., that *houses* are expected and task-relevant).

bias occurs. They also do not contain sufficient information to study the impact of task-relevance and expectation on sensory processing. Therefore, we inspected the labels given by a linear discriminant classifier trained to discriminate between face and house images. Figure 8 shows the probability that this classifier awards an ambiguous stimulus with a 'face' label, separately for each experimental condition. We again tested the difference in probability values at 172 ms after stimulus onset with RM-ANOVA, which indicated a significant main effect of expectation ($F(1, 25) = 14.84$, $p = 0.001$) and task-relevance ($F(1, 25) = 7.26$, $p = 0.012$), but no interaction effect ($F(1, 25) = 0.07$, $p = 0.794$). However, the exact probability scores given by the classifier may be the result of bias or coincidence. In order to interpret the effects of expectation and task-relevance, it may be more informative to inspect the differences between conditions.

Figure 9 shows the classification output when we only consider one factor at a time. Apart from a (non-significant) increase in the probability of a 'face' label soon after the task-relevance cue in conditions where faces were task-relevant, subtracting the averaged conditions does not reveal notable differences. Given that the behavioral effects of task-relevance and expectation on responses to ambiguous stimuli are not reflected in the labels given by a neural classifier trained on a period shortly after the onset of a face or house image (128-138 ms post stimulus onset), we may cautiously conclude that these effects are perhaps more decision-related than perceptual. For example, in case of uncertainty about the stimulus

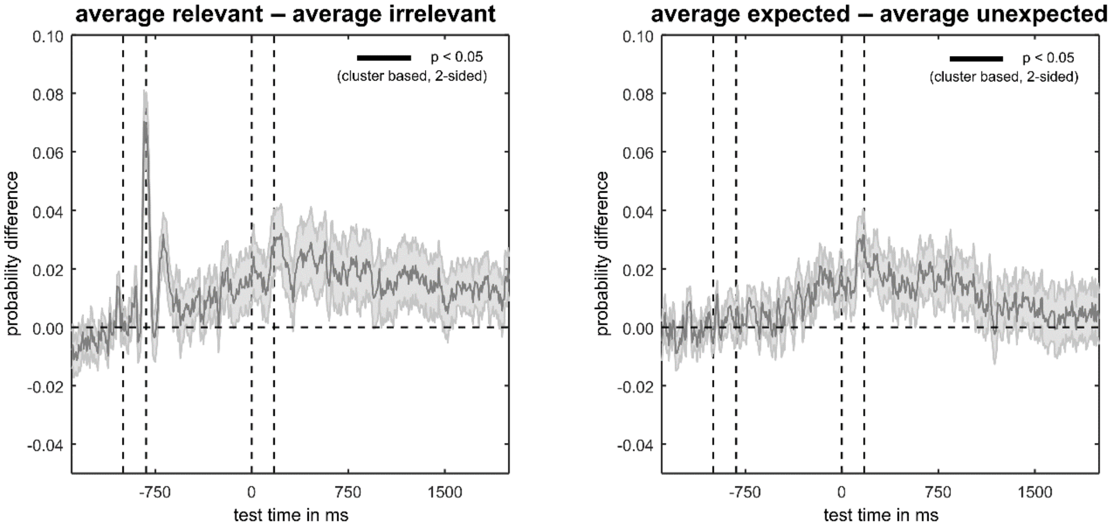


Figure 9. Main effects of task-relevance and expectation on classification of ambiguous stimuli. All scores are plotted ± SEM (light colored). Dark bold lines indicate $p < 0.05$ with cluster based permutation. **A)** shows the difference between the average probability of a 'face' label in conditions where faces were task-relevant ('expected and task-relevant' and 'unexpected and task-relevant') and the average probability of a 'face' label in conditions where faces were task-irrelevant ('expected and task-irrelevant' and 'unexpected and task-irrelevant'). **B)** shows the difference between the average probability of a 'face' label in conditions where faces were expected ('expected and task-relevant' and 'expected and task-irrelevant') and the average probability of a 'face' label in conditions where faces were unexpected ('unexpected and task-relevant' and 'unexpected and task-irrelevant').

identity, the participant may guess an appropriate answer and be biased in their guessing by the expectation and task-relevance cues.

To explore the idea that ambiguous stimuli evoke uncertainty about the stimulus identity, we repeated the classification analysis separated by response level on the PAS visibility scale. First, we investigated whether the pattern of classification varied for different levels of subjective awareness. In Figure 10 this pattern of classification (averaged across all conditions) is plotted for all PAS levels. Again, we can extract the probability values at 172 ms after stimulus onset. A RM-ANOVA (only including 15 participants with complete results for each PAS level) indicates a significant main effect of PAS level on the probability of a

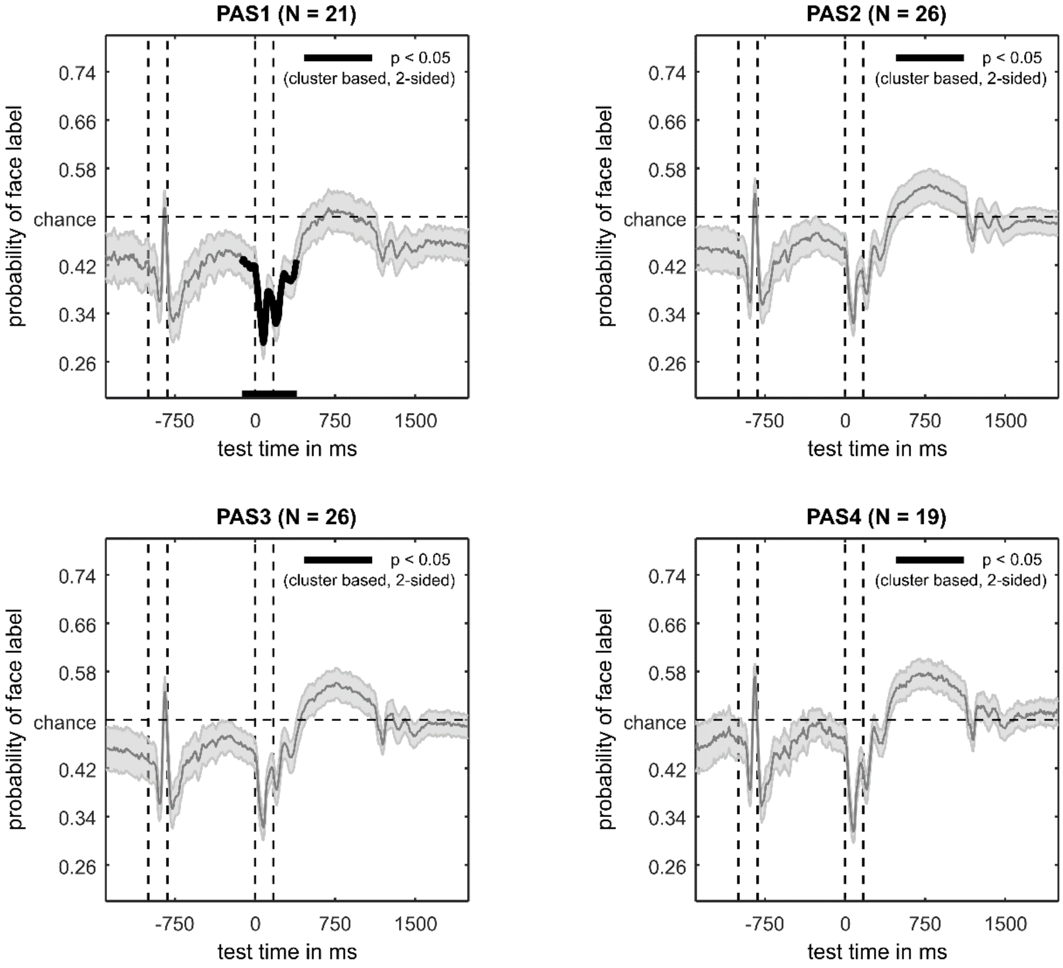


Figure 10. EEG classification labels over time for each PAS level, trained and tested on all electrodes to discriminate face and house target stimuli. Line graphs reflect the average probability of a 'face' label across participants ± SEM (light colored). Higher values indicate that the algorithm is more likely to classify the image as a face, lower values indicate that the algorithm is more likely to classify the image as a house. Dark bold lines indicate p < 0.05 with cluster based permutation. From left to right, dotted vertical lines represent cue onset (-1000 ms), peak cue decoding (-820 ms), target onset (0 ms), and peak target decoding (172 ms). Note that the number of participants of which results are averaged is not the same for each subplot. Within each level of the PAS, only participants with complete results (i.e., who responded this PAS level at least once in *each* condition) were considered for this analysis.

‘face’ label ($F(3, 12) = 3.82$, $p = 0.039$). This means that higher levels of subjective awareness are associated with an increased probability of a ‘face’ label by the neural classifier. However, it should be noted that for all PAS levels, the probability value at 172 ms after stimulus onset is lower than 0.50, which means that a ‘house’ label is more probable than a ‘face’ label. The relationship between PAS level and classification probability is not necessarily a causal one: there could be a large number of reasons behind this association. One of our suggestions is that a clearer experience of the ambiguous stimulus may counteract the general bias towards a ‘house’ label. Another possibility is that participants were more likely to report that they experienced an ambiguous stimulus clearly when they perceived it as a face. On the basis of these results, we are unable to arbitrate between these possible interpretations.

Discussion

The aim of the current research was to study the effects of expectation and task-relevance on sensory processing in conscious perception. For this purpose, we conducted an EEG experiment where participants performed a task in which they were asked to judge the orientation of task-relevant (but not of task-irrelevant) face and house stimuli, while expectation was manipulated over blocks. For each stimulus, participants also judged visibility.

Our experimental findings demonstrate a main effect of both expectation and task-relevance on the perception of task stimuli. Firstly, face and house images were more often correctly recognized when they were expected or task-relevant. Ambiguous images, containing both a face and a house image, were also more often interpreted as being of the expected or task-relevant category than of the unexpected or task-irrelevant category. Secondly, using multivariate pattern analysis of EEG signals, we were able to decode the stimulus category of expected or task-relevant stimuli more accurately than of unexpected or task-irrelevant stimuli. We did not find an interaction effect between expectation and task-relevance in our behavioral or neural measures. We further found that higher levels of subjective awareness as measured by the PAS visibility scale were associated with higher decoding accuracy, although this relationship proved unsignificant at a timepoint associated with peak decoding accuracy.

When we compare our conclusions to previously reported findings from similar studies, the absence of an interaction effect in our results is at odds with several previous studies that did find interacting effects of expectation and task-relevance (Kok et al., 2012b; Jiang et al., 2013; St. John-Saaltink et al., 2015). For example, Kok and colleagues (2012b) conducted an experiment in which participants were asked to judge the orientation of Gabor patches in cued locations, while expectations about the likely stimulus location were manipulated block-wise. They found that expected stimuli evoked stronger BOLD response in V1 when they occurred in cued (i.e., task-relevant) locations, but weaker BOLD response when they occurred in other (i.e., task-irrelevant) locations. However, in addition to the fact that there are considerable differences in task difficulty between our studies, Kok and colleagues measured stimulus-evoked activity with fMRI while we measured stimulus-evoked *patterns* of activity, focusing on early EEG signals.

Smout and colleagues (2019) propose a theory that would reconcile the contradictory findings on whether or not there is an interaction between expectation and task-relevance.

According to this theory, predictions and prediction errors are represented in distinct groups of neurons, respectively encoding feature information or mismatch information. Only the latter type of neuronal population is suggested to be affected by task-relevance. In essence, the authors pose that expectation influences the brain’s predictions (encoded by *features*), while expectation and task-relevance have interacting influences on prediction errors (encoded by *mismatch*). In the current study, we employed multivariate pattern analysis to study the patterns of brain activity associated with a certain stimulus type. As these patterns of representation are more likely reflective of feature information than mismatch information, this could possibly explain why we do not find an interacting effect of expectation and task-relevance. In agreement with this suggestion, a previous fMRI study using multivariate pattern analysis (Kok et al., 2012a) also did not find an interaction effect.

Our results could also contribute to the ongoing discussion of whether expectation impacts sensory processing at early or late stages, and the related question of whether expectation mainly impacts sensory processing or decision-making processes. While some researchers argue that prior expectations affect the earliest stages of perceptual processing by evoking (prestimulus) sensory templates (Kok et al., 2013; Kok, Failing, & de Lange, 2014; Kok, Mostert, & de Lange, 2017), others pose that expectations modulate only later (post-perceptual) stages of perceptual decision-making (Alilović et al., 2019; Bang & Rahnev, 2017; Rungratsameetaweemana et al., 2018). In the current study, we aimed to separate sensory processes from decision processes. For this reason, the linear discriminant classifier was trained on an early period (128-138 ms) of EEG data from a perceptual training task that was completely unrelated to decision-making processes in the main task, similar to the approach by Mostert and colleagues (2015). We also report *off*-diagonal decoding accuracy, meaning that the classifier is trained on one timepoint and tested on other timepoints. As the classifier performs decoding based on some discriminative neural pattern found in the training set, we can assume that accurate decoding reflects similar underlying patterns of neural activity (King & Dehaene, 2014). That is, similar to the patterns of activity evoked during early sensory processing of decision-unrelated face and house stimuli. In other words, if prior expectations were to affect sensory processing at early stages, we should see increased decoding of expected compared to unexpected stimuli in our results. While we do observe that increase to some extent, a more compelling result is that for task-relevant stimuli, compared to task-irrelevant stimuli, we observe significant prestimulus decoding of the stimulus category evoked by the task-relevance cue. In addition to a sharp peak in decoding shortly after the presentation of the task-relevance cue, we found another significant decoding cluster roughly between 500 and 100 ms before stimulus onset. This observation is similar to the prestimulus sensory templates found by Kok and colleagues (2017) which were evoked by an expectation cue. Crucially, it is precisely the use of a cue that is criticized by opponents of early expectation effects: they suggest that a cue may induce allocation of selective attention to a stimulus, causing sensory gain that would not occur due to expectation in the absence of an explicit cue (Simon, Schachtner, & Gallen, 2019). As the cue in our study was independent of the manipulation of expectation, the current results indeed suggest that it is the cue itself and not its effect on expectation that evokes the sensory templates, perhaps through the allocation of selective attention. In sum, our results support the criticism that explicit cues may conflate expectation and attention, and therefore we would like to recommend that future studies manipulate expectation (and task-relevance) without the use of explicit cues.

Finally, we would like to make a small remark about the suboptimality of our measures of perception. Firstly, both our behavioral (sensitivity) and neural (decoding accuracy) measures of stimulus judgment were optimized for comparison between conditions. For this reason, we selected trials from different blocks for the calculation of our outcome measures. In the actual task, participants would never have to decide whether they saw an image of an expected face or an expected house, as expectation was manipulated block-wise to avoid the use of an expectation cue. Thus, our ‘expected’ conditions contain face and house trials from different blocks. While this allows us to investigate, for example, whether a person’s ability to discriminate between faces and houses increases when the stimulus is expected rather than unexpected, it does come with the implicit assumption that there is no inherent difference between perceiving faces and perceiving houses. Should this assumption be violated, then this would also mean that our measures of sensitivity and decoding accuracy are biased. Therefore, our results are better interpreted in relation to each other (e.g., ‘participants were more sensitive to relevant than irrelevant stimuli’) than as exact values (e.g., ‘we can decode stimulus category with 56% accuracy’). Secondly, given that PAS levels were reported at the very end of each trial, we also recommend that results including this measure be interpreted with appropriate caution. The PAS was included with the intention to measure the subjective awareness of stimuli, but as it is responded to after judging the orientation and category of the stimulus, it may also measure some aspects of decision confidence. In addition, we did not actually manipulate visibility of the stimuli, making the PAS a purely correlational measure. Future studies could consider manipulating visibility in order to examine causal relationships.

We would like to end with several suggestions for future analyses that were outside of the scope of the current study, but that would be valuable extensions to this line of research. To begin, we have good reason to assume the involvement of multiple distinct brain areas in the effects of expectation and task-relevance on perception (see e.g., Summerfield & Egner, 2009; Summerfield & de Lange, 2014). In the current study, we considered EEG activity from all 64 electrodes in our analyses, but it would also be informative to focus on a narrower set of electrodes, such as only occipital or only frontal electrodes. Another alternative is to perform source reconstructions. This could for example be beneficial in determining whether an effect is purely sensory or more decision-related. However, since it is unlikely that the influences of expectation or task-relevance are fully localized in one area of the brain, perhaps it would be even more informative to study connectivity or communication between certain brain areas. Friston (2005) poses that perception is the result of an interaction between predictions and prediction errors, reflected by backward and forward communication between brain areas located at higher and lower levels in the cortical hierarchy. Consistent with this view, one study finds that prior expectations modulate both forward and backward connectivity between sensory and frontal regions (Rahnev, Lau, & de Lange, 2011). Given that attention is proposed to scale the size of prediction errors (Feldman & Friston, 2010; Hohwy, 2012), which are mainly propagated forward in the cortical hierarchy, one more specific research question could be whether task-relevance affects forward processing. Although several studies have tackled similar questions by examining the amplitude of the C1 ERP component (Baumgarter et al., 2017; Kelly, Gomez-Ramirez, & Foxe, 2008), results so far have been inconclusive, which could be attributed to a number of methodological pitfalls (Slagter, Alilović, & van Gaal, 2017).

A related avenue for future analyses is in the time-frequency domain. Previous research supports the possibility that information about predictions and prediction errors are signaled in distinct frequencies, where backward connections are likely associated with lower frequencies than forward connections (Bastos et al., 2015). Consistent with this view, prior knowledge or expectations have been proposed to be mainly signaled in low-frequency activity such as alpha frequency (Brodski-Guerniero et al., 2017), while prediction errors are more likely to be signaled in higher-frequency activity (Han & VanRullen, 2017). In the context of perceptual decision-making, it would be informative to investigate whether expectation mainly affects lower-frequency signals, and whether task-relevance mainly affects higher-frequency signals. Ultimately, these future studies could inform us on how the brain generates a conscious experience of the visual world.

References

Alilović, J., Timmermans, B., Reteig, L. C., van Gaal, S., & Slagter, H. A. (2019). No evidence that predictions and attention modulate the first feedforward sweep of cortical information processing. *Cerebral Cortex*, *29*(5), 2261-2278.

Bang, J. W., & Rahnev, D. (2017). Stimulus expectation alters decision criterion but not sensory signal in perceptual decision making. *Scientific Reports*, *7*(1), 1-12.

Bastos, A. M., Vezoli, J., Bosman, C. A., Schoffelen, J. M., Oostenveld, R., Dowdall, J. R., ... & Fries, P. (2015). Visual areas exert feedforward and feedback influences through distinct frequency channels. *Neuron*, *85*(2), 390-401.

Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology*, *14*(2), 143-177.

Bowman, H., Filetti, M., Wyble, B., & Olivers, C. (2013). Attention is more than prediction precision [Commentary on target article]. *Behavioral and Brain Sciences*, *36*(3), 206-208.

Bradley, A. P. (1997). The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition*, *30*(7), 1145-1159.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*(4), 433-436.

Brodski-Guerniero, A., Paasch, G. F., Wollstadt, P., Özdemir, I., Lizier, J. T., & Wibral, M. (2017). Information-theoretic evidence for predictive coding in the face-processing system. *Journal of Neuroscience*, *37*(34), 8273-8283.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181-204.

Davenport, J. L., & Potter, M. C. (2004). Scene consistency in object and background perception. *Psychological Science*, *15*(8), 559-564.

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9-21.

Egner, T., Monti, J. M., & Summerfield, C. (2010). Expectation and surprise determine neural population responses in the ventral visual stream. *Journal of Neuroscience*, *30*(49), 16601-16608.

Endl, W., Walla, P., Lindinger, G., Lalouschek, W., Barth, F. G., Deecke, L., & Lang, W. (1998). Early cortical activation indicates preparation for retrieval of memory for faces: An event-related potential study. *Neuroscience Letters*, *240*(1), 58-60.

Fahrenfort, J. J., van Driel, J., van Gaal, S., & Olivers, C. N. L. (2018). From ERPs to MVPA Using the Amsterdam Decoding and Modeling Toolbox (ADAM). *Frontiers in Neuroscience*, 12.

Feldman, H., & Friston, K. (2010). Attention, uncertainty, and free-energy. *Frontiers in Human Neuroscience*, *4*, 215.

Friston, K. (2003). Learning and inference in the brain. *Neural Networks*, *16*(9), 1325-1352.

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1456), 815-836.

Friston, K. (2008). Hierarchical models in the brain. *PLoS Computational Biology*, *4*(11), e1000211.

Friston, K. (2009). The free-energy principle: A rough guide to the brain? *Trends in Cognitive Sciences*, *13*(7), 293-301.

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, *11*(2), 127-138.

Gordon, N., Tsuchiya, N., Koenig-Robert, R., & Hohwy, J. (2019). Expectation and attention increase the integration of top-down and bottom-up signals in perception through different pathways. *PLoS Biology*, *17*(4), e3000233.

Han, B., & VanRullen, R. (2017). The rhythms of predictive coding? Pre-stimulus phase modulates the influence of shape perception on luminance judgments. *Scientific Reports*, *7*(1), 1-10.

Hohwy, J. (2012). Attention and conscious perception in the hypothesis testing brain. *Frontiers in Psychology*, *3*, 96.

Jiang, J., Summerfield, C., & Egner, T. (2013). Attention sharpens the distinction between expected and unexpected percepts in the visual brain. *Journal of Neuroscience*, *33*(47), 18438-18447.

Kanade, T., Cohn, J. F., & Tian, Y. (2000, March). Comprehensive database for facial expression analysis. In *Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580)* (pp. 46-53). IEEE.

Kelly, S. P., Gomez-Ramirez, M., & Foxe, J. J. (2008). Spatial attention modulates initial afferent activity in human primary visual cortex. *Cerebral Cortex*, *18*(11), 2629-2636.

King, J. R., & Dehaene, S. (2014). Characterizing the dynamics of mental representations: The temporal generalization method. *Trends in Cognitive Sciences*, *18*(4), 203-210.

Kok, P., Brouwer, G. J., van Gerven, M. A., & de Lange, F. P. (2013). Prior expectations bias sensory representations in visual cortex. *Journal of Neuroscience*, *33*(41), 16275-16284.

Kok, P., Failing, M. F., & de Lange, F. P. (2014). Prior expectations evoke stimulus templates in the primary visual cortex. *Journal of Cognitive Neuroscience*, *26*(7), 1546-1554.

Kok, P., Jehee, J. F., & de Lange, F. P. (2012a). Less is more: Expectation sharpens representations in the primary visual cortex. *Neuron*, *75*(2), 265-270.

Kok, P., Mostert, P., & de Lange, F. P. (2017). Prior expectations induce prestimulus sensory templates. *Proceedings of the National Academy of Sciences*, *114*(39), 10473-10478.

Kok, P., Rahnev, D., Jehee, J. F., Lau, H. C., & de Lange, F. P. (2012b). Attention reverses the effect of prediction in silencing sensory signals. *Cerebral Cortex*, *22*(9), 2197-2206.

Meijs, E. L., Mostert, P., Slagter, H. A., de Lange, F. P., & van Gaal, S. (2019). Exploring the role of expectations and stimulus relevance on stimulus-specific neural representations and conscious report. *Neuroscience of Consciousness*, *2019*(1), niz011.

Meijs, E. L., Slagter, H. A., de Lange, F. P., & van Gaal, S. (2018). Dynamic interactions between top–down expectations and conscious awareness. *Journal of Neuroscience*, *38*(9), 2318-2327.

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.

Mostert, P., Kok, P., & de Lange, F. P. (2015). Dissociating sensory from decision processes in human perceptual decision making. *Scientific Reports*, *5*(1), 1-13.

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

Rahnev, D., Lau, H., & De Lange, F. P. (2011). Prior expectation modulates the interaction between sensory and prefrontal regions in the human brain. *Journal of Neuroscience*, *31*(29), 10741-10748.

Ramsøy, T. Z., & Overgaard, M. (2004). Introspection and subliminal perception. *Phenomenology and the Cognitive Sciences*, *3*(1), 1-23.

Rungratsameetaweemana, N., Itthipuripat, S., Salazar, A., & Serences, J. T. (2018). Expectations do not alter early sensory processing during perceptual decision-making. *Journal of Neuroscience*, *38*(24), 5632-5648.

Sherwell, C., Garrido, M. I., & Cunnington, R. (2017). Timing in predictive coding: The roles of task relevance and global probability. *Journal of Cognitive Neuroscience*, *29*(5), 780-792.

Slagter, H. A., Alilovic, J., & Van Gaal, S. (2018). How early does attention modulate visual information processing? The importance of experimental protocol and data analysis approach. *Cognitive Neuroscience*, *9*(1-2), 26-28.

Smout, C. A., Tang, M. F., Garrido, M. I., & Mattingley, J. B. (2019). Attention promotes the neural encoding of prediction errors. *PLoS Biology, 17*(2), e2006812.

St. John-Saaltink, E., Utzerath, C., Kok, P., Lau, H. C., & De Lange, F. P. (2015). Expectation suppression in early visual cortex depends on task set. *PLoS One*, *10*(6), e0131172.

Summerfield, C., & de Lange, F. P. (2014). Expectation in perceptual decision making: neural and computational mechanisms. *Nature Reviews Neuroscience*, *15*(11), 745-756.

Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends in Cognitive Sciences*, *13*(9), 403-409.

Summerfield, C., & Koechlin, E. (2008). A neural representation of prior information during perceptual inference. *Neuron*, *59*(2), 336-347.

Summerfield, C., Trittschuh, E. H., Monti, J. M., Mesulam, M. M., & Egner, T. (2008). Neural repetition suppression reflects fulfilled perceptual expectations. *Nature Neuroscience, 11*(9), 1004.

van Driel, J., Olivers, C. N., & Fahrenfort, J. J. (2021). High-pass filtering artifacts in multivariate classification of neural time series data. *Journal of Neuroscience Methods*, *352*, 109080.

Weaver, M. D., Fahrenfort, J. J., Belopolsky, A., & van Gaal, S. (2019). Independent Neural Activity Patterns for Sensory- and Confidence-Based Information Maintenance during Category-Selective Visual Processing. *eNeuro*, *6*(1), ENEURO.0268-18.2018.

Williams, N. R., Willenbockel, V., & Gauthier, I. (2009). Sensitivity to spatial frequency and orientation content is not specific to face perception. *Vision Research*, *49*(19), 2353-2362.

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*, *109*(9), 3593-3598.