Real-world expectations and their affective value modulate object processing

Laurent Caplette[1], Frédéric Gosselin[1], Martial Mermillod[2] et Bruno Wicker[1,3]

[1]Département de psychologie, Université de Montréal, Montréal, Québec, H3C 3J7, Canada

[2]LPNC, CNRS & Université Grenoble Alpes, 38058 Grenoble, France

[3]LNC, CNRS & Aix-Marseille Université, 13331 Marseille, France

Correspondence should be addressed to:

Bruno Wicker

LNC, Aix-Marseille Université

3, Place Victor-Hugo

13331 Marseille cedex 3, France

Email : bruno.wicker@univ-amu.fr

Phone : +33(0)413550900

Fax : +33(0)413550958

Abstract

It is well known that expectations influence how we perceive the world. Yet the neural mechanisms underlying this process remain unclear. Studies have focused so far on artificial contingencies between simple neutral cues and events. Real-world expectations are however often generated from complex associations between potentially affective contexts and objects learned over a lifetime. In this study, we used fMRI to investigate how object processing is influenced by neutral and affective context-based expectations. First, we show that the precuneus and the inferotemporal cortex are more active during object recognition when expectations have been elicited a priori, irrespectively of their validity or their affective intensity. This result supports previous hypotheses according to which these brain areas integrate contextual expectations with object sensory information. Early visual areas, on the contrary, are more active during object recognition when no prior expectation has been elicited by a context. Finally, we show that BOLD activity is enhanced in early visual areas when objects are less expected, but only when contexts are neutral; the reverse effect is observed when contexts are affective. This result supports a recent proposal that affect modulates predictions in the brain. Together, our results help elucidate the neural mechanisms of real-world expectations.

Keywords: prediction, context, fMRI, object recognition, emotion.

We expect to find hairdryers in bathrooms, tombstones in cemeteries, and baguettes in bakeries, but more rarely tombstones in bathrooms, refrigerators in cemeteries and hairdryers in bakeries. That is, we live in a world where most objects are associated with specific contexts. Throughout a lifetime of experiences, we come to learn these associations, which lead us to form expectations about the objects to be encountered when we navigate the world. Such expectations about specific objects can lead to their enhanced recognition once they appear in our visual field (see Henderson and Hollingsworth 1999; Chun 2000; Bar 2004, for reviews). Indeed, we tend to recognize more easily objects if they are preceded by a congruent context than by an incongruent one (Palmer 1975).

Multiple authors have proposed that the brain acts as a predictive machine and attempts to match incoming sensory inputs with top-down expectations (e.g., von Helmholtz 1866; Gregory 1980; Ullman 1995; Friston 2005; Grossberg 2009; de Lange et al. in press). According to the "predictive coding" theory, predictions from high-level areas are fed back to adjacent lower-level areas and compared with incoming sensory signals, such that only the discrepancy between these two signals — the prediction error — and not the stimulus *per se*, is transmitted up the visual hierarchy (Friston 2005; see also Mumford 1992; Ullman 1995; Rao and Ballard 1999). Following the recognition of a context (e.g., a kitchen), high-level areas could thus implicitly compute the probabilities of different objects being seen — the predictions — and gradually feed them back to lower and lower visual areas: sensory neurons representing features of probable objects (e.g., a knife) would thus fire before any object enters the visual field. When an object finally stimulates photoreceptors, only the prediction error travels back to higher-level areas for the object's recognition (Friston 2005). In support of this model, prior expectation of a given visual stimulus has been found to elicit a pattern of activity specific to that stimulus in the primary visual cortex (Kok et al. 2014, 2017; Hindy et al. 2016). It was also reported that valid prior expectations result in reduced neural activity in sensory cortices (Summerfield et al.

2008; den Ouden et al. 2010; Alink et al. 2010; Kok et al. 2012a; Todorovic and de Lange 2012).

Some predictions may require a different mechanism than feedback from adjacent visual areas (Hindy et al. 2016). For example, Bar and colleagues (Bar 2003, 2004; Bar et al. 2006) proposed that context-based expectations could be preferentially initiated in the parahippocampal (PHC) and retrosplenial (RSC) cortices, using the rapidly extracted coarse visual information, and then directly fed back to the inferior temporal cortex and fusiform gyrus where they would be compared and integrated with sensory object information. Although this hypothesis has not been fully tested, highly contextual objects do activate the PHC and RSC (Bar and Aminoff 2003), and contextually related objects do activate the precuneus, which is part of the RSC, more than contextually unrelated objects (Livne and Bar 2016). Moreover, when object and context are viewed together, the context seems to facilitate neural representations of objects in object-selective regions, and this facilitation correlates with activity in the PHC and RSC (Brandman and Peelen 2017). The hippocampus also seems to play an important role in the generation of expectations (Hindy et al. 2016; Kok and Turk-Browne in press).

Interestingly, an unexpected stimulus will result in an enhanced sensory response even when stimuli are task-irrelevant (Alink et al. 2010; den Ouden et al. 2010; Yaron et al. 2012; Kok et al. 2012a) and even when the cues are predictive of the targets before but not during the experiment (i.e. when there is an equal number of “predictive” and “non-predictive” trials; Hindy et al. 2016). This suggests that contingencies engraved in our memory are not easily ignored, even when we know that they are not helpful for the task at hand (see Summerfield and Egner 2009; Kok et al. 2012b), and that predictions are automatic to some extent.

Most studies examining the effect of predictions on perception have used very simple cues such as tones (Summerfield and Koechlin 2008; den Ouden et al. 2010; Kok et al. 2012a,

2017) or a repetition of the same object (Summerfield et al. 2008; Todorovic and de Lange 2012). By contrast, expectations about everyday objects usually stem from other objects or the surrounding context. It is likely that predictions based on these more complex stimuli originate mostly from high-level areas such as the PHC or even the prefrontal cortex (Bar 2004). Relatedly, the effect of predictions has not been considered in the setting of an ecological object recognition task. Simple detection tasks (Jiang et al. 2013), delayed discrimination tasks (Kok et al. 2012a, 2014, 2017) or categorization tasks using few alternatives (den Ouden et al. 2010; Kok et al. 2012b) are typically used. Moreover, previous studies on prediction have manipulated predictability by artificial means, either by repeating and alternating stimuli (Summerfield et al. 2008; Todorovic and de Lange 2012), by having stimuli appearing after different cues with different probabilities during the experiment (den Ouden et al. 2010; Kok et al. 2012a, 2012b, 2014, 2017; Jiang et al. 2013), or by developing arbitrary contingencies over a short time scale before the experiment (Hindy et al. 2016). Associations between contexts and objects formed over a lifetime of experiences may involve mechanisms distinct from these, not only because they might be more deeply engrained in our brains, but also because they are more complex, with all objects being associated with all contexts with varying probabilities; for example, memories developed a long time ago are typically more independent of the hippocampus than newly formed memories (Squire et al. 2004).

Real-world expectations are often tinted by some affective value. A visual context can elicit emotional reactions that may influence the recognition of objects in the scene. For example, in an emotional context (e.g., a cemetery), the affective value may be partially processed before the scene's objects (e.g., a tombstone) and in fact help predict their identities (Barrett and Bar 2009). Moreover, the affective value of the context may prioritize the processing of specific behaviorally relevant features and objects and influence the speed and efficiency of recognition. As such, expectations activated by a specific context might be

fundamentally different whether the context is emotional or not; and this is likely true even for small valences (Lebrecht et al. 2012). Most of these propositions have not yet been investigated, but related interactions between affect and associations have been reported in a few studies. For example, it has been reported that our affective state influences the breadth of the associations we make (Storbeck & Clore 2005). In addition, contextual and affective values of everyday objects interact in their influence over the information used to recognize them rapidly (Caplette et al. 2014). Relatedly, associative and affective processing both seem to take place in the same brain area, the medial orbitofrontal cortex, hinting to a link between these processes (Shenhav et al. 2013).

In the present study, we addressed all these shortcomings. We investigated how realistic object recognition mechanisms are influenced by task-irrelevant high-level expectations generated by a predictive or non-predictive visual context. More specifically, subjects were shown visual scenes followed by briefly presented images of objects while being scanned in an MRI scanner. Different pairings of target objects and scenes were used to produce three main contextual conditions: predictive (e.g., kitchen and refrigerator), non-predictive (e.g., bathroom and refrigerator), and no-context (e.g., scrambled image and refrigerator). To ensure that the subjects recognized the object, an object name was shown afterwards and subjects had to determine if the name matched the object or not (Figure 1). The use of unambiguous everyday objects and scenes allowed us to use associations between objects and contexts formed over a lifetime of experiences. Furthermore, subjects did not benefit from taking contexts into account while doing the task since scenes were followed by an unrelated object as often as they were followed by a related object (Hindy et al. 2016). Hence, evidences of predictions in this experiment would imply a fast and automatic mechanism rather than a more explicit and conscious one. To uncover potential brain areas where object sensory information and contextual expectations are compared and integrated, we first investigated whether object

processing was influenced by expectations, independently of their validity. We then examined which areas preferentially represent matches and prediction errors by comparing valid and invalid expectations. Finally, we sought to identify whether the effects of valid and invalid predictions were the same in the case of contexts with affective value.

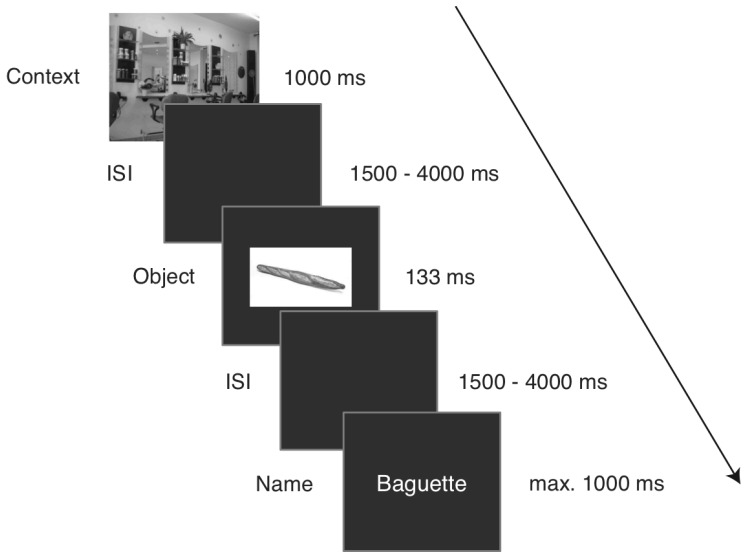

Figure 1. Example of a trial (Non-Predictive Neutral condition). The object image and name have been enlarged for better viewing. Scene and object images were in color during the experiment.

Materials and Methods

Participants

Seventeen healthy adults (9 female; mean age = 24.8; SD = 4.3) were recruited on the campus of Aix-Marseille Université. Participants did not suffer from any neurological, psychological or psychiatric disorder and were free of medication. The experimental protocol was approved by the ethics board of CPP Sud-Méditerranée 1 and the study was carried in accordance with the approved guidelines. Written informed consent was obtained from all participants after the procedure had been fully explained, and a monetary compensation was provided upon completion of the experiment.

Stimuli

Thirty-three public domain scene color images were selected from the internet as context images. In a first validation study, an independent sample of 22 subjects identified what they thought the images represented, indicated if the context elicited an emotion and, if so, what were the valence (negative to positive, from 0 to 10) and intensity (no emotion to very intense emotion, from 0 to 10). Following this study, 32 visual scenes were selected and split in Affective (e.g., cemetery, beach, luxury hotel) and Neutral (e.g., swimming pool, airport, kitchen) categories at the median of the intensity scores (5.19).

In a second validation study, 35 different subjects were shown the context names and had to give the names of three objects with a high probability of being present in that context. Ninety-six color images of objects (three for each context) corresponding to the most cited names were selected for the experiment (e.g., swimsuit, diving board and pool ladder for swimming pool; see table S1 for all contexts and related objects). For each context, the experimenters also chose three non-associated objects (selected from the ones that had never been associated with the context in the second validation study). Every object was the associated object of only one context and the non-associated object of only one other context; moreover, for each context, each one of the three non-associated objects was associated with a different context.

Finally, we randomized the phases of the mean of the context images in the Fourier domain – separately for each RGB color channel – to obtain 96 different phase-scrambled images.

Experimental Design

The LabVIEW (National Instruments Inc., Austin, TX, USA) software was used to project stimuli during the experiment. Stimuli were projected to a screen positioned in the back of the scanner using a video projector. Subjects could see the video reflected in a mirror (15 x 9 cm) suspended 10 cm in front of their face and subtending visual angles of 42 degrees horizontally and 32 degrees vertically.

Each trial was built as follows: a large cue image (see below) spanning the whole screen during 1 s, a black screen during 1.5 to 4 s (duration randomly selected from a truncated exponential distribution with mean of 2 s), a centered object image on a black background during 133 ms, a black screen during 1.5 to 4 s, and an object name on a black background shown until the subject answered or for a maximum of 1 s (Figure 1). Subjects answered by pressing one of two buttons on a hand-held response device to indicate if the name corresponded to the object, which occurred on 80% of the trials. A black screen was displayed for an additional 1 s between trials.

On a third of the trials (Predictive condition), the cue image was a scene associated with the object following it (e.g., an airport and a suitcase); on another third (Non-Predictive condition), it was a scene not associated with the object following it (e.g., a church and a tennis racket); on the final third (No-Context condition), it was a scrambled image (always a different one; see *Stimuli* section). Each object was shown once in each of these conditions, for a total of 288 trials. Furthermore, Predictive and Non-Predictive conditions were each split evenly into the Affective and Neutral subconditions, following the affective intensity of the context. There was therefore a total of 5 conditions: Predictive Affective (or Pred-Aff for short), Predictive Neutral (Pred-Neut), Non-Predictive Affective (noPred-Aff), Non-Predictive Neutral (noPred-Neut) and No-Context (noCont).

The order of trials was randomized. Randomized trials were divided in 3 fixed functional data acquisition runs of 96 trials. Each functional run lasted between 10 and 12 mins, with short breaks between them. The order of the 3 runs was counterbalanced across subjects.

Data acquisition

Functional imaging data were acquired with an ADVANCE 3 Tesla scanner (Bruker Inc., Ettlingen, Germany) equipped with a 2-channel head-coil. Functional images sensitive to BOLD contrast were acquired with a T2*-weighted gradient echo EPI sequence (TR 2400 ms, TE 30 ms, matrix 64 x 64 mm, FOV 192 mm, flip angle 81.6°). Thirty-six slices with a slice gap of 0 mm were acquired within the TR; voxels were 3 x 3 x 3 mm. Between 303 and 311 volumes were acquired in each run, excluding the six dummy scans acquired at the beginning of each run for signal stabilization. Additionally, a high resolution (1 x 1 x 1 mm) structural scan was acquired from each participant with a T1-weighted MPRAGE sequence.

Preprocessing and statistical analysis

For preprocessing and statistical analysis, the SPM8 software (http://www.fil.ion.ucl.ac.uk/spm/), running in the MATLAB environment (Mathworks Inc., Natick, MA, USA), was used. T1-weighted structural images were segmented into white matter, gray matter and cerebrospinal fluid, and warped into MNI space. Functional images were realigned, unwarped and corrected for geometric distortions using the field map of each participant, slice time corrected, coregistered to the structural image of the corresponding participant, and smoothed using a 6 mm FWHM isotropic Gaussian kernel.

Two different GLM analyses were performed in each subject. In a first context-and-object model, one event was modelled per trial with onset time at the beginning of the presentation of the context and offset time at the moment of the subject's response. Each

condition of interest (Pred-Aff, Pred-Neut, noPred-Aff and noPred-Neut) and the noCont condition were modelled by convolving box-car functions with SPM8's canonical hemodynamic response function; the six motion parameters were also included in the model. This model was used to investigate the main effect of prediction (Pred-Aff + Pred-Neut – noPred-Aff – noPred-Neut), the main effect of affective value (Pred-Aff – Pred-Neut + noPred-Aff – noPred-Neut), and the interaction between predictive and affective values (Pred-Aff – Pred-Neut – noPred-Aff + noPred-Neut). It was possible to use the complete trials for these comparisons because the objects and contexts are the same across the conditions and only the associations between them differ; resulting activations are necessarily correlated to object recognition since the associations become clear to the subject only when the object is shown. The use of the complete trials was preferred whenever possible to increase signal-to-noise ratio.

In a second object-only model, two events were modelled on each trial. Object events (the regressors of interest) were modelled with onset time at the beginning of the presentation of the object and offset time at the moment of the subject's response; each condition (Pred-Aff, Pred-Neut, noPred-Aff, noPred-Neut and noCont) was modelled separately. Scene events (regressors of no interest) were modelled with onset time at the beginning of the presentation of the context and offset time at the beginning of the presentation of the object. These regressors were all modelled by convolving box-car functions with SPM8's canonical hemodynamic response function; the six motion parameters were also included in the model. Since visual scenes are different in the noCont condition, we used this model where the events of interest did not include the scene to compare the noCont condition with other conditions.

A temporal high-pass filter (cut-off of 128 s) was used to remove low-frequency drifts, and temporal autocorrelation across scans was modelled with an AR(1) process. Contrasts were then computed at the subject level and used for group analyses using one-sample *t*-tests. All voxels inside the brain were analyzed; we maintained the familywise error rate of $p < .05$, two-

tailed, at the cluster level (primary threshold of $p < .001$, uncorrected) using random field theory (Friston et al. 1994). The Anatomy (Eickhoff et al. 2005) and WFU-PickAtlas (Maldjian et al. 2003) toolboxes were used to identify activated brain regions based on peak Montreal Neurological Institute (MNI) coordinates.

Results

Behavioral results

Mean accuracy was 97.1% (σ = 2.8%) for the Pred-Neut condition, 97.2% (σ = 2.1%) for the Pred-Aff condition, 96.9% (σ = 2.6%) for the noPred-Neut condition, 95.0% (σ = 2.5%) for the noPred-Aff condition and 96.6% (σ = 2.5%) for the noCont condition. When comparing Pred, noPred and noCont together, there was no effect of condition on accuracy ($F(2,16) = 2.32$, $p = .11$). When comparing all conditions except noCont together, there was a significant main effect of predictive value ($F(1,16) = 10.36$, $p = .005$), a marginally significant main effect of affective value ($F(1,16) = 4.42$, $p = .051$), and a marginally significant interaction between affective and predictive values ($F(1,16) = 3.97$, $p = .063$).

Mean response time was 640 ms (σ = 121 ms) for the Pred-Neut condition, 650 ms (σ = 120 ms) for the Pred-Aff condition, 638 ms (σ = 110 ms) for the noPred-Neut condition, 636 ms (σ = 127 ms) for the noPred-Aff condition and 632 ms (σ = 116 ms) for the noCont condition. When comparing Pred, noPred and noCont together, there was no effect of condition on response time ($F(2,16) = 1.53$, $p = .23$). When comparing all conditions except noCont together, there was no main effect of affective or predictive value and no interaction (Fs(1,16) = 1.67, 0.25 and 0.64 respectively, ps > .20).

fMRI results

To investigate the potential effect of the generation of explicit contextual expectations (occurring only in the Pred and noPred conditions) on brain activity, we contrasted the Pred and noPred conditions with the noCont condition using the object-only model. Four clusters were significantly more activated in the Pred and noPred conditions than in the noCont condition (Figure 2a; Table 1): one cluster extending from the left middle temporal gyrus to the parietal cortex (angular gyrus), one in the right middle temporal gyrus, one in the right angular gyrus, and one in the precuneus. The reverse contrast revealed the specific activation of a large bilateral cluster extending from the middle occipital gyrus to the anterior fusiform, and a smaller cluster in the right calcarine gyrus (Figure 2a; Table 1).

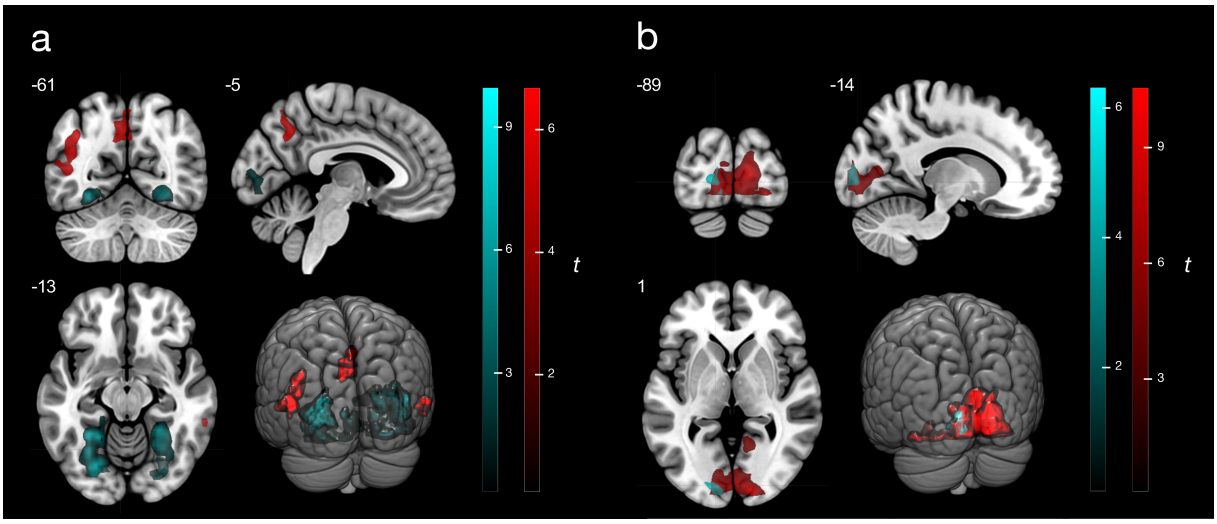

Figure 2. A) Significant clusters for the noCont > (Pred + noPred) (in red) and the (Pred + noPred) > noCont (in cyan) contrasts. **B)** Significant clusters for the Pred-Aff > noPred-Aff (in red) and the noPred-Neut > Pred-Neut (in cyan) contrasts.

Using the context-and-object model, we then investigated whether there was a main effect of predictive value (Pred vs noPred), a main effect of the context's affective intensity (Aff vs Neut), and an interaction between predictive and affective values on brain areas involved in object recognition. There were no significant main effects of predictive and affective values. However, there was a significant interaction between predictive and affective values for a large bilateral cluster extending from the calcarine sulci to the cuneus (Table 1). When decomposing the interaction into simple effects, these areas, along with the right fusiform

gyrus and, in another cluster, the hippocampus (subiculum, CA3) and the parahippocampal gyrus, were significantly more active in the Pred-Aff condition than in the noPred-Aff condition (Figure 2b; Table 1). No cluster was significantly more active in the noPred-Aff condition than in the Pred-Aff condition. When looking at the neutral conditions, no cluster was significantly more active in the Pred-Neut condition than in the noPred-Neut condition, but the left lingual and superior occipital gyri were more active in the noPred-Neut condition than in the Pred-Neut condition (Figure 2b; Table 1). Interestingly, voxels at the intersection of the calcarine sulci, lingual gyrus and superior occipital gyrus were both part of a cluster that was more active in the Pred-Aff condition compared to the noPred-Aff condition and part of a cluster that was more active in the noPred-Neut condition compared to the Pred-Neut condition.

Table 1. Montreal Neurological Institute (MNI) coordinates and T values for significantly activated brain regions.

Brain regions	Peak MNI coordinates			Nb of voxels	Peak T value
	x	y	z		
(Pred + noPred) > noCont					
L Angular	-42	-54	30	238	6.67
L Middle Temporal	-54	-48	6		5.89
L Precuneus	0	-60	39	101	5.93
R Middle Temporal	63	-42	-3	63	6.51
R Angular	39	-54	33	36	5.96
noCont > (Pred + noPred)					
R Middle Occipital	36	-81	6	1607	9.97
R Lingual	27	-45	-9		9.97
L Inferior Occipital Gyrus	-27	-81	-12		8.72
L Lingual Gyrus	-24	-48	-9		8.01
L Fusiform Gyrus	-30	-63	-9		8.01
R Calcarine	15	-51	9	59	5.58
Predict. x Affect.					
R Cuneus	15	-96	6	799	10.45
L Superior Occipital	-15	-96	9		8.17
L Calcarine	-9	-87	-6		8.16
Pred-Aff > noPred-Aff					
R Calcarine	15	-93	6	1047	10.56
L Calcarine	-3	-84	-6		7.87
R Fusiform	27	-84	-12		7.44
R Hippocampus Subiculum	27	-21	-15	44	5.61
R Parahippocampal	30	-27	-21		5.05
noPred-Neut > Pred-Neut					
L Lingual	-12	-90	0	57	6.31
L Superior Occipital	-12	-93	12		5.37

Discussion

Our first aim was to investigate how the generation of expectations about objects from a preceding context might modulate the activity of brain areas involved in object perception. To our knowledge, differential object processing depending on the presence of prior expectations, independently of their degree of validity, had never been investigated before. We found significantly more activation in the precuneus, the middle temporal gyri and the parietal cortex, when contextual expectations were generated prior to object perception, suggesting that these high-level areas are chiefly associated with object processing when expectations are generated. These activations specifically represent an interaction between contextual expectations and object bottom-up sensory information: activity related solely to object processing is cancelled out because the objects are the same in both conditions, and activity related solely to the prior presentation of the context is regressed out in the GLM.

According to a popular hypothesis, top-down contextual predictions would be combined with bottom-up sensory information to facilitate object recognition in the inferotemporal cortex (Bar 2004). The observed interaction between object and context processing in the middle temporal gyrus (a part of the inferotemporal cortex) supports this hypothesis. The precuneus and the parietal cortex, which are also activated in this contrast, have previously been linked to episodic memory retrieval and contextual associative processing (Lundstrom et al. 2005; Aminoff et al. 2007; Livne and Bar 2016; Brandman and Peelen 2017) which both require the integration of stored representations with incoming sensory information. Moreover, the precuneus of an observer that views several objects simultaneously is more activated when these objects are contextually related than when they are not (Livne and Bar 2016); this suggests that the contextual representations elicited by some of these objects are compared to other objects. Recently, activity in the retrosplenial complex, a region comprising the precuneus, has been shown to correlate with supra-additive decoding of objects embedded in scenes,

suggesting that the area is responsible for a scene-based facilitation of object representations (Brandman and Peelen 2017). Interestingly, the interaction we observed between context and object information in the precuneus is also supra-additive (i.e. there is a remaining positive activation after considering the main effects of object and context). We extend previous results by showing that the precuneus integrates object sensory information with valid or invalid scene-based expectations generated prior to object presentation.

The reverse contrast, associated with visual processing of objects when no expectation (neither valid nor invalid) had been generated from a context, yielded bilateral activation of lateral occipital, fusiform and lingual areas. This result indicates that these early and intermediate visual areas are predominantly associated with processing of sensory information specifically when there are no expectations, or at least no expectations generated by a context. Activated voxels may be part of areas primarily associated with the processing of sensory information shared by the vast majority of objects (e.g., intermediate spatial frequencies; Caplette et al. 2014), which is thus reduced when almost any object is expected.

We then investigated whether there was an effect of prediction error or match, i.e. whether some areas were more active at the presentation of the object when the object followed a predictive context or when the object followed a non-predictive context. Looking first at neutral contexts, we observed that the lingual and superior occipital gyri were more active in the non-predictive condition than in the predictive condition. This pattern of activations echoes results obtained in many studies in which activations in sensory areas when the prediction is incorrect are seen as markers of prediction error (e.g., Summerfield & Koechlin 2008; Kok et al. 2012a, 2012b). Notably, most, if not all, of these studies also used affectively neutral cues. Combining neutral and affective contexts, we did not observe any significant difference between predictive and non-predictive conditions. A significant interaction between emotional and predictive values confirmed that the affective intensity of the prediction influenced

prediction error mechanisms. This interaction was visible in a large bilateral cluster in low-level occipital areas comprising the lingual and superior occipital gyri and the calcarine sulci. The calcarine sulci showed the opposite pattern of activations when the context was emotional: they were more active when the object was predicted than when it was not.

Our results are not compatible with the proposal that a subject's internal affective state is altering the content of their predictions about object identities (Barrett and Bar 2009). According to this idea, the affective value of a preceding context (or even a simultaneous context or the object itself; see Barrett and Bar 2009) would alter the subject's bodily state and bring additional information that could be used by the brain to predict the identity of perceived objects. Consequently, a similar pattern of results should be visible for neutral and emotional contexts, with only a greater difference in activation between predicted and unpredicted objects for emotional contexts than for neutral contexts (due to the additional emotional information).

Our results are compatible, however, with the general idea that affect interacts with predictive processing (Barrett and Simmons 2015; Miller and Clark 2018). One possibility recently put forward by some authors is that, rather than contributing to the content of the predictions, a subject's internal affective state modulates the precision of the predictions (Miller and Clark 2018). In recent formulations of predictive coding (Feldman and Friston 2010), the prediction error is weighted by the reliability, or precision, of sensory information. When precision is low, prediction errors are down-weighted and observers rely more on predictions; when it is high, prediction errors are up-weighted and observers rely more on the sensory input. Because this weighting only occurs for neurons representing prediction error and not for neurons representing predictions (Friston 2009; Kok et al. 2012b), it should lead to an interaction between prediction error and precision (Rao 2005; Friston 2009; Kok et al. 2012b) exactly like the one we observed between predictive and affective values. This weighting may be performed by the pulvinar (Saalman et al. 2012; Kanai et al. 2015; Miller and Clark 2018),

which is believed to bias visual processing toward emotionally salient signals (Adolphs 2002; de Gelder et al. 2011; Pessoa 2013).

Kok and colleagues (Kok et al. 2012b) reported a similar reversal of the prediction error effect in the early visual cortex for task relevant stimuli. They argued that this effect was caused by endogenous attention enhancing the precision of the predictions (Rao 2005; Feldman & Friston 2010). This cannot be the cause of the effect we observed however, since attention was not manipulated in our study and our stimuli were all similarly task relevant. Furthermore, reaction times and accuracies were similar for affective and neutral contexts; in fact, there even was a tendency toward lower rather than higher recognition rates for affective contexts (p = .051; see also Caplette et al. 2014). This indicates that exogenous attention was also similar between the conditions and implies that the prediction error reversal in our study was not caused by an increase in attention.

Relatedly, the same interaction applied to behavioral recognition rates also approached statistical significance (p = .06), suggesting that differences in brain activation may be translated into behavioral differences. Contrary to what was found at the brain level, however, a main effect of predictive value (irrespectively of the context's affective value) attained statistical significance at the behavioral level (see also Mermillod et al. 2018). This could be due to opposite or different activations in some brain areas leading to similar behavioral effects, to the behavioral effects being caused by the interactive action of a collection of brain regions, or simply to a lack of statistical power or resolution to uncover any such effect in the brain.

We also observed significantly more activation in the medial temporal lobe, specifically in the hippocampus and parahippocampal cortex, when an object was following a predictive emotional context compared to when it was following a non-predictive emotional context. Higher-level areas have been shown to display enhanced activation when there is a match between expectation and sensory information (Murray et al. 2002; Harrison et al. 2007;

Summerfield et al. 2006; Summerfield & Koechlin 2008); this is expected in the predictive coding framework because higher-level areas should "complete" sensory information in lower areas with the predictions they represent (Summerfield & Koechlin 2008). Moreover, the parahippocampal cortex is known to represent contextual information (Epstein & Kanwisher 1998; Bar & Aminoff 2003; Aminoff et al. 2013) and has been hypothesized to initiate predictions based on contextual information (Bar 2004). The hippocampus has also been shown to be the source of top-down predictions (Davachi & DuBrow 2015; Hindy et al. 2016; Kok and Turk-Browne in press). Here, we show that the parahippocampal cortex is more activated when the object matches context-based expectations, providing the clearest evidence to date that it does initiate predictions based on the context. We further show that the affective value of expectations seems to have an impact on the enhanced match response observed in the medial temporal lobe. It has long been known that affective memories tend to be more consolidated (e.g. McGaugh and Roozendaal 2002; McGaugh 2004), and more generally that the amygdala and the hippocampal complex interact in several ways (see Phelphs 2004 for a review), so this raises the intriguing possibility that the initiation of predictions in these areas is modulated by affective content.

In summary, we uncovered that real-world expectations initiated by contexts, irrespectively of their degree of validity, lead to the specific activation of high-level visual and associative brain areas during subsequent object recognition and to a relative deactivation of some of the areas classically involved in bottom-up object perception. To our knowledge, the influence of the presence of expectations on object processing had not been investigated prior to this study. Furthermore, the context's affective value interacted with the validity of the prediction it had initiated: classical prediction error effects were only observed with neutral contexts, and a complete reversal of these effects was observed when contexts were emotional. This result is not compatible with the idea that the affective value of a stimulus, and the ensuing

internal bodily state of the subject, are contributing to the creation of predictions; but it is compatible with a modulatory role of affective value over the weight of predictions in perception (Barrett and Bar 2009; Miller and Clark 2018). Finally, hippocampal and parahippocampal areas were selective to matches between affective contexts and objects, suggesting that contextual predictions are initiated in these areas (Bar 2004; Kok and Turk-Browne in press) and that this process is modulated by affective value.

Our results are in line with the brain being proactive and constantly forming predictions about future sensory inputs. Besides predictive coding theory, several other theories of perceptual inference are articulated around this idea (Mumford 1994; Lee and Mumford 2003; Grossberg 2013; Fiser et al. 2010; Friston 2010; Orbán et al. 2016; see de Lange et al. in press) and our results are also compatible with them. In conclusion, these findings bring us closer to uncovering the neural mechanisms by which we integrate prior knowledge with sensory inputs to optimize perception.

Funding

This study was funded by the Fondation Planiol (BW), by the Institut Universitaire de France (MM) and by the Social Sciences and Humanities Research Council of Canada (Joseph-Armand-Bombardier Scholarship and Michael-Smith Foreign Study Supplement to LC).

References

Adolphs R. 2002. Neural systems for recognizing emotion. Curr Opin Neurobiol. 12(2):169–177.

Alink A, Schwiedrzik CM, Kohler A, Singer W, Muckli L. 2010. Stimulus predictability reduces responses in primary visual cortex. J Neurosci. 30(8):2960–2966.

Aminoff EM, Gronau N, Bar M. 2007. The parahippocampal cortex mediates spatial and nonspatial associations. Cereb Cortex. 17(7):1493–1503.

Aminoff EM, Kveraga K, Bar M. 2013. The role of the parahippocampal cortex in cognition. Trends Cogn Sci. 17(8): 379–390.

Bar M. 2003. A cortical mechanism for triggering top-down facilitation in visual object recognition. J Cogn Neurosci. 15(4):600–609.

Bar M. 2004. Visual objects in context. Nat Rev Neurosci. 5(8):617–629.

Bar M, Aminoff E. 2003. Cortical analysis of visual context. Neuron. 38(2):347–358.

Bar M, Kassam KS, Ghuman AS, Boshyan J, Schmid AM, Dale AM, Hämäläinen MS, Marinkovic K, Schacter DL, Rosen BR, et al. 2006. Top-down facilitation of visual recognition. Proc Natl Acad Sci U S A, 103(2):449–454.

Barrett LF, Bar M. 2009. See it with feeling: affective predictions during object perception. Philos Trans R Soc Lond B Biol Sci. 364(1521):1325–1334.

Barrett LF, Simmons WK. 2015. Interoceptive predictions in the brian. Nat Rev Neurosci. 16:419-429.

Brandman T, Peelen MV. 2017. Interaction between scene and object processing revealed by human fMRI and MEG decoding. J Neurosci. 37(32):7700–7710.

Caplette L, West G, Gomot M, Gosselin F, Wicker B. 2014. Affective and contextual values modulate spatial frequency use in object recognition. Front Psychol. 5:512.

Chun MM. 2000. Contextual cueing of visual attention. Trends Cogn Sci. 4(5):170–178.

Davachi L, DuBrow S. 2015. How the hippocampus preserves order: the role of prediction and context. Trends Cogn Sci. 19(2):92–99.

de Gelder B, van Honk J, Tamietto M. 2011. Emotion in the brain: of low roads, high roads and roads less travelled. Nat Rev Neurosci. 12(7):425.

de Lange FP, Heilbron M, Kok P. In press. How do expectations shape perception? Trends Cogn Sci.

den Ouden HEM, Daunizeau J, Roiser J, Friston KJ, Stephan KE. 2010. Striatal prediction error modulates cortical coupling. J Neurosci. 30(9):3210–3219.

Eickhoff SB, Stephan KE, Mohlberg H, Grefkes C, Fink GR, Amunts K, Zilles K. 2005. A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. Neuroimage. 25(4):1325–1335.

Epstein R, Kanwisher N. 1998. A cortical representation of the local visual environment. Nature. 392:598–601.

Feldman H, Friston KJ. 2010. Attention, uncertainty, and free-energy. Front Hum Neurosci. 4:215.

Fiser J, Berkes P, Orbán G, Lengyel M. 2010. Statistically optimal perception and lerning: from behavior to neural representations. Trends Cogn Sci. 14(3):119–130.

Friston KJ. 2005. A theory of cortical responses. Philos Trans R Soc Lond B Biol Sci. 360(1456):815–836.

Friston KJ. 2009. The free-energy principle: a rough guide to the brain? Trends Cogn Sci. 13(7):293–301.

Friston KJ. 2010. The free-energy principle: a unified brain theory? Nat Rev Neurosci. 11(2):607–608.

Friston KJ, Worsley KJ, Frackowiak RSJ. 1994. Assessing the significance of focal activations using their spatial extent. Hum Brain Mapp. 1:214-220.

Gregory RL. 1980. Perceptions as hypotheses. Phil Trans R Soc Lond B. 290:181–197.

Grossberg S. 2009. Cortical and subcortical predictive dynamics and learning during perception, cognition, emotion and action. Philos Trans R Soc Lond B Biol Sci, 364(1521):1223–1234.

Grossberg S. 2013. Adaptive resonance theory: how a brain learns to consciously attend, learn, and recognize a changing world. Neural Networks. 37:1–47.

Harrison LM, Stephan KE, Rees G, Friston KJ. 2007. Extra-classical receptive field effects measured in striate cortex with fMRI. Neuroimage. 34(3):1199–1208.

Henderson JM, Hollingworth A. 1999. High-level scene perception. Annu Rev Psychol, 50:243–271.

Hindy NC, Ng FY, Turk-Browne NB. 2016. Linking pattern completion in the hippocampus to predictive coding in visual cortex. Nat Neurosci. 19(5):665–667.

Jiang J, Summerfield C, Egner T. 2013. Attention sharpens the distinction between expected and unexpected percepts in the visual brain. J Neurosci. 33(47):18438–18447.

Kanai R, Komura Y, Shipp S, Friston KJ. 2015. Cerebral hierarchies: predictive processing, precision and the pulvinar. Philos Trans R Soc Lond B Biol Sci. 370(1668):69–81.

Kok P, Failing MF, de Lange FP. 2014. Prior expectations evoke stimulus templates in the primary visual cortex. J Cogn Neurosci. 26(7):1546–1554.

Kok P, Jehee JFM, de Lange FP. 2012. Less is more: expectation sharpens representations in the primary visual cortex. Neuron. 75(2) :265–270.

Kok P, Mostert P, de Lange FP. 2017. Prior expectations induce prestimulus sensory templates. Proc Natl Acad Sci U S A. 114(39):10473–10478.

Kok P, Rahnev D, Jehee JFM, Lau HC, de Lange, FP. 2011. Attention reverses the effect of prediction in silencing sensory signals. Cereb Cortex. 22(9):2197–2206.

Kok P, Turk-Browne NB. In press. Stimulus prediction in the hippocampus. J Neurosci.

Lebrecht S, Bar M, Barrett LF, Tarr MJ. 2012. Micro-valences: perceiving affective valence in everyday objects. Front Psychol. 3:107.

Lee TS, Mumford D. 2003. Hierarchical Bayesian inference in the visual cortex. J Opt Soc Am A Opt Image Sci Vis. 20(7):1434–1448.

Livne T, Bar M. 2016. Cortical integration of contextual information across objects. J Cogn Neurosci. 28(7):948–958.

Lundstrom BN, Ingvar M, Petersson KM. 2005. The role of preuneus and left inferior frontal cortex during source memory episodic retrieval. Neuroimage. 27(4):824–834.

Maldjian JA, Laurienti PJ, Kraft RA, Burdette JH. 2003. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. Neuroimage. 19(3):1233–1239.

McGaugh JL. 2004. The amygdala modulates the consolidation of memories of emotionally arousing experiences. Annu Rev Neurosci. 27:1–28.

McGaugh JL, Roozendaal B. 2002. Role of adrenal stress hormones in forming lasting memories in the brain. Curr Opin Neurobiol. 12(2):205–210.

Mermillod M, Grynberg D, Pio-Lopez L, Rychlowska M, Beffara B, Harquel S, Vermeulen N, Niedenthal PM, Dutheil F, Droit-Volet S. 2018. Evidence of rapid modulation by social information of subjective, physiological, and neural responses to emotional expressions. Front Behav Neurosci. 11:231.

Miller M, Clark A. 2018. Happily entangled: prediction, emotion, and the embodied mind. Synthese. 195(6):2559–2575.

Mumford D. 1992. On the computational architecture of the neocortex: II. The role of cortico-cortical loops. Biol Cybern. 66:241.

Mumford D. 1994. Neuronal architectures for pattern-theoretic problems. In: Koch C, Davis JL, editors. Large-scale neuronal theories of the brain. Cambridge (MA): MIT press.

Murray SO, Kersten D, Olshausen BA, Schrater P, Woods DL. 2002. Shape perception reduces activity in human primary visual cortex. Proc Natl Acad Sci U S A. 99(23):15164–15169.

Orbán G, Berkes P, Fiser J, Lengyel M. 2016. Neural variability and sampling-based probabilistic representations in the visual cortex. Neuron. 92(2):530–543.

Palmer SE. 1975. The effects of contextual scenes on the identification of objects. Mem Cogn. 3(5):519–526.

Pessoa L. 2013. The cognitive-emotional brain: from interactions to integration. Cambridge (MA): MIT press.

Phelps EA. 2004. Human emotion and memory: interactions of the amygdala and hippocampal complex. Curr Opin Neurobiol. 14(2):198–202.

Rao RPN. 2005. Bayesian inference and attentional modulation in the visual cortex. Neuroreport. 16(16):1843–1848.

Rao RPN, Ballard DH. 1999. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nat Neurosci. 2(1):79–87.

Saalman YB, Pinsk MA, Wang L, Li X, Kastner S. 2012. The pulvinar regulates information transmission between cortical areas based on attention demands. Science. 337(6095):753–756.

Shenhav A, Barrett LF, Bar M. 2013. Affective value and associative processing share a cortical substrate. Cogn Affect Behav Neurosci. 13(1):46–59.

Squire LR, Stark CEL, Clark RE. 2004. The medial temporal lobe. Annu Rev Neurosci. 27:279–306.

Storbeck J, Clore GL. 2005. With sadness comes accuracy; with happiness, false memory: mood and the false memory effect. Psychol Sci. 16(10) :785–791.

Summerfield C, Egner T. 2009. Expectation (and attention) in visual cognition. Trends Cogn Sci. 13(9):403-409.

Summerfield C, Egner T, Greene M, Koechlin E, Mangels J, Hirsch J. 2006. Predictive codes for forthcoming perception in the frontal cortex. Science. 314(5803):1311–1314.

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

Summerfield C, Trittschuh EH, Monti JM, Mesulam MM, Egner T. 2008. Neural repetition suppression reflects fulfilled perceptual expectations. Nat Neurosci. 11(9):1004–1006.