NeuroImage 173 (2018) 351–360

Contents lists available at ScienceDirect

NeuroImage

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

Distributed and opposing effects of incidental learning in the human brain

Michelle G. Hall [a,*], Claire K. Naughtin [a], Jason B. Mattingley [a,b], Paul E. Dux [a]

[a] School of Psychology, The University of Queensland, QLD, 4072, Australia
[b] Queensland Brain Institute, The University of Queensland, QLD, 4072, Australia

A R T I C L E I N F O

Keywords:
Statistical learning
Expectation
fMRI
Multi-voxel pattern analysis
Visual cognition

A B S T R A C T

Incidental learning affords a behavioural advantage when sensory information matches regularities that have previously been encountered. Previous studies have taken a focused approach by probing the involvement of specific candidate brain regions underlying incidentally acquired memory representations, as well as expectation effects on early sensory representations. Here, we investigated the broader extent of the brain's sensitivity to violations and fulfilments of expectations, using an incidental learning paradigm in which the contingencies between target locations and target identities were manipulated without participants' overt knowledge. Multivariate analysis of functional magnetic resonance imaging data was applied to compare the consistency of neural activity for visual events that the contingency manipulation rendered likely versus unlikely. We observed widespread sensitivity to expectations across frontal, temporal, occipital, and sub-cortical areas. These activation clusters showed distinct response profiles, such that some regions displayed more reliable activation patterns under fulfilled expectations, whereas others showed more reliable patterns when expectations were violated. These findings reveal that expectations affect multiple stages of information processing during visual decision making, rather than early sensory processing stages alone.

Introduction

Despite being dizzyingly complex and often unpredictable, the world we live in is highly structured in space and time. The human visual system is well equipped to capitalise on this structure through mechanisms involved in incidental learning (Perruchet and Pacton, 2006; Turk-Browne, 2012). For example, mention of a jackhammer elicits a strong sense of context, conjuring mental images of cement and hard hats and other objects that have consistently been encountered in relation to one another. Intuitively, we know that it is easier to react to something predictable, in context, than to something surprising. Behavioural research has confirmed this, showing that participants are able to recognise temporal sequences and spatial configurations of objects that they have encountered previously (Chun and Jiang, 1998; Fiser and Aslin, 2001, 2002; Hall et al., 2015; Turk-Browne et al., 2005), and are faster to identify objects if they appear in a predictable context (Oliva and Torralba, 2007; Turk-Browne et al., 2005; Turk-Browne et al., 2010). This adaptation to predictable visual input often proceeds incidentally (i.e., without instruction, but not necessarily outside of awareness), as is the case with much contextual learning in real world scenarios (Bar, 2004).

The consequences of incidental learning for behaviour are well documented (for review, see Turk-Browne, 2012), but how this is borne out in the brain is less clear. Using functional magnetic resonance imaging (fMRI), Turk-Browne et al. (2009) reported activation in medial temporal and striatal memory systems in response to structured sequences of shapes. Given that these regions are heavily implicated in both explicit and implicit memory processes (Gabrieli, 1998; Poldrack et al., 2001), they are key candidate regions where statistical regularities might be encoded in a durable form. The role of medial temporal structures in particular has been implicated in a number of subsequent studies (Schapiro et al., 2014; Schapiro et al., 2012; Turk-Browne et al., 2010). In addition to activity in these memory regions, Turk-Browne and colleagues' (2009) work revealed anticipatory activity in visual processing regions. This highlights a potential mechanism by which learned regularities come to influence behaviour: in response to learned sensory cues, associations are retrieved from memory, and modulate activity in sensory processing regions in anticipation of the upcoming stimulus.

In line with this possibility, theories of predictive coding propose that learned “priors” bias sensory processing (e.g., Rao and Ballard, 1999). Studies using explicit cueing paradigms have demonstrated that expectations affect fMRI responses in visual processing regions before the

* Corresponding author. School of Psychology, The University of Queensland, McElwain Building, St Lucia, QLD, 4072, Australia
E-mail address: michelle.hall@uqconnect.edu.au (M.G. Hall).

https://doi.org/10.1016/j.neuroimage.2018.02.068
Received 7 December 2017; Received in revised form 27 February 2018; Accepted 28 February 2018
Available online 30 March 2018

appearance of a stimulus, as evidenced by elevated activity in brain regions that encode features of the expected stimulus (Esterman and Yantis, 2010), and anticipatory activity patterns that reflect stimulus-specific representations (Kok et al., 2014). Responses evoked by the onset of a stimulus are also contingent on expectations, with surprising events robustly evoking enhanced activity, relative to expected events, in visual processing regions ("expectation suppression"; Amado et al., 2016; Egner et al., 2010; Summerfield et al., 2008). Kok et al., 2012 reported that elevated activity for unexpected versus expected orientation grating stimuli was coupled with reduced sensitivity of multivariate decoding. Specifically, activity evoked by surprising stimuli did not reflect enhanced stimulus representation, but instead likely reflected a combination of top-down predictions, bottom-up sensory processing, and prediction error signals, aggregating to a noisy population response. Conversely, activity associated with expected stimuli likely reflected integration of top-down predictions and congruent bottom-up sensory signals, and therefore minimal prediction error. In sum, processing of expected stimuli in sensory areas is associated with a sharpened and efficient population response (Grill-Spector et al., 2006; Kersten et al., 2004).

Although these focal perspectives investigating expectations and incidental learning in specific brain regions are critical for understanding the neural mechanisms underlying expectation processing, they do not provide a comprehensive account of the consequences of incidental learning across multiple processing stages underlying perceptual decision making (Behrmann and Plaut, 2013). Beyond anticipatory responses in early sensory regions, incidental learning is likely to influence a range of downstream cognitive processes such as attentional orienting, response selection, motor control, and other executive functions (Heekeren et al., 2008; Mesulam, 1998), suggesting additional neural loci whose representations may be affected by expectations. For example, whereas fulfilled expectations can facilitate efficient behavioural responses, violated expectations are likely to trigger a range of cognitive control processes to overcome the perceptual conflict, such as attentional recruitment and inhibition of an anticipated response. Similarly, in executive control regions, expectations might modulate evidence accumulation, task-set maintenance and updating, or adaptive control (Dosenbach et al., 2007), which may manifest as sharper activity patterns under violated expectations, when greater control is required.

To assess the broad range of brain regions that might be influenced by expectations, we analysed the consistency of fMRI activity patterns. This implementation of multi-voxel pattern analysis (MVPA) assumes that reliable voxel-by-voxel patterns of BOLD activity reflect the representation carried in the region (Haxby et al., 2001; Kriegeskorte et al., 2006; Kriegeskorte et al., 2008). Using this approach to study episodic memory, Xue et al. (2010) reported that the fidelity of activation patterns across a range of visual processing regions was associated with better memory encoding and retrieval. Extending this MVPA approach to representational content beyond sensory and memory domains, Esterman et al. (2009) reported reliable activation patterns underlying cognitive control states in parietal cortex, specifically the reconfiguration of task set underlying shifts of attention, reconfiguration of working memory representations, and shifts of categorization rules. This work suggests that reliable indices of cognitive control states can also be extracted from higher-level brain regions using MVPA. Similarly, Garner and Dux (2015) showed reliable divergence between multi-voxel patterns underlying distinct task sets across frontal, parietal, and subcortical regions.

Here we investigated how incidental learning affects neural representations underlying perceptual decision-making judgments across the whole brain, using behaviour and fMRI. We first conducted a behavioural experiment to validate a novel incidental learning paradigm. In a second experiment, we used fMRI to identify brain regions that showed reliable differences in the stability of patterns of neural activity, depending on whether visual stimuli were congruent or incongruent with incidentally learned expectations. We found a distributed set of cortical and subcortical regions that displayed sensitivity to incidentally learned visual expectations. These regions showed opposing response profiles, such that patterns of activity in some regions were more consistent under fulfilled expectations, while those for a number of regions were more consistent under violated expectations. This work highlights the multifaceted nature of statistical learning on visual processing in the human brain.

Experiment 1 (behaviour)

Materials and methods

Experimental design

To evaluate how expectations resulting from incidental learning affect activity throughout the brain, we developed a behavioural task in which the contingencies between four fractal-like targets and four target locations – one in each of the four quadrants of the visual field – were manipulated (Fig. 1). Each target location was associated with two possible target shapes – one shape that appeared there frequently and one shape that appeared there infrequently. The same targets were associated with a second location, but with the opposite contingency mapping. The remaining two target locations were associated with a second pair of targets, again with the contingencies exchanged across the two locations. Each trial had a single target that was thus either likely at the cued location (Frequent target condition), or unlikely at the cued location (Infrequent target condition).

Importantly, the contingency manipulation was implemented such that each shape was presented an equal number of times over the course of each run (at two locations, across trials with Frequent and Infrequent targets), controlling for potential effects of overall stimulus frequency. Moreover, the target location was spatially cued prior to target onset on each trial. This ensured that spatial attention was matched for trials with Frequent and Infrequent targets at target onset, and allowed expectations to develop in response to the location cue. Participants were not informed of the location-shape contingencies, and were simply instructed to respond to the identity of the target as quickly and as accurately as possible on each trial. Although participants may have become aware of the contingencies over the course of the experiment, any expectations, whether outside of awareness or not, were acquired incidentally. For completeness, we assessed participants' awareness of the contingency manipulation with a brief questionnaire at the end of the experiment. Finally, the paradigm allowed for expectation effects to be measured over the course of the experiment by comparing response times to Frequent targets and Infrequent targets. This is in contrast to other incidental learning paradigms in which learning is often probed offline in a separate test phase (e.g., Fiser and Aslin, 2001).

Participants

Twenty-two participants were recruited from The University of Queensland community and were paid $10 for participation. This target sample size was specified prior to data collection, based on previous studies investigating incidental learning (e.g., Fiser and Aslin, 2001), as well behavioural piloting. Two participants were excluded for poor performance (>3 sd. above the group mean response time, or >10% missing responses), and one participant was excluded for colour-blindness. Analysis included data from the remaining 19 participants (mean age = 22.26y [s.d. 5.24y], 5 male). Participants had normal or corrected-to-normal vision, and provided written informed consent in accordance with a protocol approved by The University of Queensland ethics committee.

Stimuli and procedure

In Experiment 1, we validated the incidental learning task (see also Hall et al., in press), and explored how different contingencies between locations and target stimuli affected target identification.

Stimuli were fractal-like shapes ($2^\circ \times 2^\circ$ of visual angle; Fig. 1), which were generated using custom MATLAB (MathWorks, Natick, USA) code, based on Miyashita et al. (1991). Each participant was allocated a unique