Donnelly, Gemma (2022) Top-down amplification of predicted visual input behind a frosted occluder. PhD thesis.

https://theses.gla.ac.uk/82909/

Top-down Amplification of Predicted Visual Input behind a Frosted Occluder

Gemma Donnelly, BSc (Hons.), MSc

Submitted in fulfilment of the requirements for the Degree of Doctor of Philosophy

Institute of Neuroscience & Psychology, College of Medical & Veterinary Life Sciences, University of Glasgow

May 2022

Contains Studies performed in:
Centre for Cognitive Neuroimaging
Institute of Neuroscience &
Psychology
University of Glasgow
G12 8QB
United Kingdom

Abstract

This thesis is comprised of five chapters. It includes two experimental chapters in which I detail both psychophysical and fMRI studies carried out at the University of Glasgow as part of this PhD project. These are followed by a literature review which outlines the implementation of ultra-high-resolution fMRI, both generally within the field and within a specific project proposal.

Chapter 1 is a general introduction. I outline the broad organisation and basic functions of the visual system at the pre-cortical and cortical stages, in turn. I then discuss the concept of feedback within the visual system, outlining what feedback is, what it does and how it is implemented before outlining the rationale for the thesis.

Chapter 2 is an experimental chapter detailing a series of psychophysical experiments. These experiments employ a partial occlusion paradigm to explore how top-down predicted information can influence the processing of degraded feedforward input. Throughout the experimental series, different aspects of this question are addressed in order to investigate whether the consistency of contextual information influences the detection and/or recognition of low-contrast visual scenes.

Chapter 3 is another experimental chapter which details two 3T fMRI experiments. These projects also employed a partial occlusion paradigm to investigate contextual modulation on the processing of degraded feedforward input at the neuronal level in early visual cortex. Both univariate and multivariate analysis techniques were used to reveal the impact of consistency within top-down information.

Chapter 4 contains a literature review which looks into ultra-high-resolution fMRI. Here, I detail the motivation behind the development of higher resolution imaging as well as potential confounds and limitations. I also outline adaptations required at higher resolution in terms of data acquisition and analysis as well as briefly exploring layer-specific findings within the visual cortex. Finally, I propose a 7T fMRI project that would continue to explore the influence of top-

down predictions on the processing of degraded visual input by expanding the investigation to a laminar level.

Chapter 5 is a general discussion which summarises the key points from each of the previous chapters and briefly discusses their conceptual relation to the current field and beyond.

Table of Contents

List of Tables

List of Figures

Publications

Brown, H. D. H., Gouws, A. D., Vernon, R. J. W., Lawrence, S. J. D., Donnelly, G., Gill, L., Gale, R. P., Baseler, H. A., & Morland, A. B. (2021). Assessing functional reorganization in visual cortex with simulated retinal lesions. *Brain Structure and Function*, *226*(9), 2855–2867.

Oral Presentations

G Donnelly, J Bergmann, T Lux, M Bennett, L.S Petro, W.A Phillips, & L Muckli (2019). *Predictive processing helps categorisation by amplifying responses to low contrast stimuli (fMRI of V1)*. Society for Neuroscience. Chicago, Illinois, 19-23 October 2019.

Posters

G Donnelly, J Bergmann, M Bennett, L.S Petro, & L Muckli (2018). *The influence of top down predictions on the processing of low contrast feedforward input.* Workshop on Predictive Processing. Donostia/San Sebastian, Spain, 20-22 June 2018.

G Donnelly, J Bergmann, M Bennett, L.S Petro, & L Muckli (2018). *The influence of top down contextual predictions on the processing of degraded feedforward information.* British Association for Cognitive Neuroscience, Annual Scientific Meeting. Glasgow, United Kingdom, 6th-7th September 2018.

G Donnelly, J Bergmann, M Bennett, L.S Petro, & L Muckli (2018). *The influence of top down contextual predictions on the processing of low contrast feedforward input*. European Conference of Visual Perception. Trieste, Italy, 26-30 August 2018.

G Donnelly, T Lux, J Bergmann, M Bennett, & L Muckli (2019). *The influence of context on behavioural and BOLD responses to low contrast information.* European Conference of Visual Perception. Leuven, Belgium, 25-29 August 2019.

G Donnelly, J Bergmann, T Lux, M Bennett, L.S Petro, W.A Phillips & L Muckli (2020). *Investigating predictive processing of low-contrast stimuli using psychophysics and fMRI.* Federation of European Neuroscience Societies. Virtual conference, 11-15 July 2020.

Attended (without presentation)

Human Brain Project Summit (2017). Glasgow, United Kingdom, 17-20 October 2017.

Organization for Human Brain Mapping (2019). Rome, Italy, 9-13 June 2019.

Acknowledgement

I would like to start by thanking my supervisor; Professor Lars Muckli who encouraged me to apply for the PhD, provided support throughout the entire process and has given me opportunities to attend conferences and present my work, making the whole experience beyond what I even hoped for. Secondly, I would like to thank Dr Lucy Petro who has given me much needed help and guidance throughout the past 4 years and has always been on hand to listen and offer advice. I really would not have been able to do this PhD without either of them.

I also want to thank my extended supervision team; Edwin Robertson, Christoph Scheepers and Martin Lages – although contact has been brief, they have always made a point of being there for me if I needed any extra guidance. A special thanks goes to Frances whose patience in teaching me how to scan is really appreciated. Also, to Bill Phillips whose ideas inspired the project and advice throughout has been a real help.

I feel very lucky to have worked in the lab that I have, I have made some great friends through the Muckli lab and hope to continue visiting Paesano with them long after my PhD is finished. In particular, I'd like to thank Johanna, Matt, Tyler, Tatjana and Clement who have all helped me with the workings of the projects or thesis at some stage (and Johanna's inspirational quotes have really kept me going). Then Michele, Yulia, Angus, Susan, Pauliina and Bianca who have all shared tips to overcome the many problems encountered within the PhD. All of the above have made working in the lab a great experience.

I could not have got through the past few years without a good group of friends. Luckily, I found some who were going through the PhD process at the same time, so I owe a huge thanks to Heather, Katy, Kirstin, Caitlin, Lauren and Gabby. They have all been invaluable support not only with their PhD advice but have also been there for countless days/nights out and trips away. I'm looking forward to celebrating all of our achievements as soon as we can. Also, thanks to my friends at home who have been there for much needed distractions, dinners, and drinks whenever I've gone back to visit.

My flatmate Kaseem also deserves a special mention, she has had to put up with hours of me panicking in some way about work-related problems but has also become a lifelong friend and I don't think there is anyone else that would have put up with living with me during the past few years.

A big thank you to Charlie who has helped me in so many ways through the most stressful few months and of course my family, especially my parents, who have unexpectedly had me move back in with them for the duration of the write-up, (of which I'm sure I have been a delight to live with) but have always been there for me with advice when I've needed it most, as always. I could not have done this PhD without them, so for that, I owe them a lot.

I would also like to thank my viva examiners Stephanie Rossit, Alessio Fracasso and convenor Guillaume Rousselet for making the viva experience an enjoyable one and for providing useful discussions and feedback - which I hope I have done justice by incorporating into the final version of this thesis.

And finally, a last minute thank you to John Shaw, Zirui Zhang and Yingying Huang for being so willing to help me overcome the boundaries of remote working when working on my corrections.

Author's Declaration

I declare that, except where explicit reference is made to the contribution of others, this thesis is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution.

Gemma Donnelly

Definitions/Abbreviations

ANOVA	Analysis of variance
BAC	Backpropagation-activated coupling
BBR	Boundary-based registration
CR	Correct rejection (Signal Detection Theory)
EPI	Echo-planar imaging
FA	Falser alarm (Signal Detection Theory)
fMRI	Functional magnetic resonance imaging
FOV	Field of view
GE	Gradient-echo
GLM	General linear model
GLMM	General linear mixed model
H	Hit (Signal Detection Theory)
HDR(F)	Haemodynamic response (function)
HSF	High spatial frequency
ISI	Inter-stimulus-interval
LGN	Lateral geniculate nucleus
LSF	Low spatial frequency
M	Miss (Signal Detection Theory)
MVPA	Multivariate pattern analysis
PVC	Primary visual cortex
RF	Receptive field

ROI	Region of interest
SD	Standard deviation
SDT	Signal Detection Theory
SE	Standard error
SVM	Support vector machine
TE	Time to echo
TR	Repetition time
UHF	Ultra-high field
UHR	Ultra-high resolution
V1	Primary visual cortex
V2	Visual area 2 (early visual cortex)
VA	Visual angle
VOI	Volume of interest

1 General Introduction

1.1 Organisation and properties of the Visual System

As one of our most salient senses, vision allows us to effectively perceive and navigate our environment. Along with other sensory processes, vision guides our action and behaviour, allowing us to interact with the external world. Outside of visual neuroscience, the focus is often on the optics of the eyes and the retina (Wässle, 2004) when thinking about vision, however, the majority of visual processing occurs in the cortex, resulting in a large portion of the cortex being devoted to this sensory modality (Zhaoping, 2014). An understanding of the cortical components of vision is therefore crucial in gaining an insight into vision as a whole.

1.1.1 Pre-Cortical Processing

Vision begins with the eyes. The retinae in the eyes detect the available visual information which the brain processes. Each eye detects input from the entire visual field. Light pigments (photons) in the visual environment travel through the various layers of the eye until they reach the retina, located at the back of the eye. Comprised of five layers, the retina contains photoreceptors which are responsible for capturing information about the available light entering the eyes (Goebel et al., 2012). Originally thought of as a pre-filter to visual processing (Gollisch & Meister, 2010), the role of the retina is now believed to be more prolific, even deemed a peripheral component of the brain (Dowling, 1987), rather than a sensory organ (Goebel et al., 2012). The principal task of the retina is to convert light into neural signals in a process known as transduction (Berry et al., 1999) which allows for subsequent cortical processing.

There are two main types of photoreceptors: rods and cones (Boycott & Wässle, 1999). Rods are the more abundant of the two and are adapted to detect photons in low-light environments. They are situated in the periphery of the retina and are absent in the centre, meaning peripheral vision is a function reserved for rods. Cones on the other hand, are responsible for colour vision and do not function in dim light levels . Cones occupy the central retina, in an area known as the fovea, where visual acuity is highest (Goebel et al., 2012). Visual acuity decreases with eccentricity, due to the depletion of cone receptors toward the periphery (Zhaoping, 2014). Cones are preferentially sensitive to specific wavelengths of visible light. Humans are trichromats; meaning we have three types of cone photoreceptors, allowing a broad spectrum of visible light to be processed, and giving rise to the many colours we can observe. Short (blue), medium (green) and long (red) wavelengths are preferentially captured by S, M and L cone receptors, respectively (Wässle, 2004) - a function dictated by the absorption capabilities of the visual pigments within each cone type.

Signals from the photoreceptors are passed on via a series of excitatory/inhibitory inputs from several retinal interneuron sub-types before reaching the retinal ganglion cells, which are considered the retinal output cells. Retinal ganglion cells have receptive field properties which allow parallel processing streams pertaining to different features of the visual signal to be formed. This is an important element of visual processing which extends into the cortex. The receptive fields of retinal ganglion cells are comprised of a centre and surround region which work antagonistically to portray the visual signal efficiently. Such antagonistic firing mechanisms mean that the ganglion neurons will fire most rapidly when the centre and surround regions are differentially activated, such that the centre and surround are not simultaneously stimulated. Having an on/off system such as this holds several advantages in terms of effective processing. Contours are more readily detected as changes in luminance across the visual scene are highlighted, as opposed to overall net luminance. As a whole, the collective receptive fields of the retinal ganglion cells cover the entire visual field, with each neuron evaluating a specific (and overlapping) portion, depending on the location and size of its centre and surround. The output of retinal ganglion cells compresses the visual signal by an order of magnitude to be transmitted via their axons to the brain for further processing. This bundle of retinal ganglion cell axons is more commonly referred to as the optic nerve. The portion of retina connected to the optic nerve contains no photoreceptors, leaving a blind spot in the visual field. However, this goes unnoticed perceptually due to cortical ‘filling in’. See (Zhaoping, 2014) for further details on retinal ganglion cells.

The optic nerve transports the visual signal from the eye to the brain and is the point of the visual system in which the visual signal is converged to the lowest number of neurons (Berry et al., 1999). The optic nerve from each eye traverses posteriorly until they conjoin in the optic chiasm. Here, a portion of the axons from each optic nerve crosses hemisphere, resulting in two optic tracts which contain visual representations of the contralateral visual field, such that visual cortex in each hemisphere receives input solely from the contralateral visual field (Goebel et al., 2012).

An important processing stage prior to cortical involvement is the thalamus. The optic tracts terminate mainly in the Lateral Geniculate Nucleus (LGN), a portion of the thalamus dedicated to early visual processing (Dagnelie, 2011; Goebel et al., 2012). The distinct neural pathways that arise in the retina pertaining to the processing of different aspects of the visual input, innervate different layers of the LGN. Commonly, these pathways are termed the Magnocellular (M) and Parvocellular (P) pathways and are each primarily concerned with the processing of different features of the visual signal (Tobimatsu et al., 1995). Neurons in the Magnocellular pathway are more responsive to information regarding movement and contrast, whereas Parvocellular neurons are more sensitive to wavelength properties, and thus, colour vision. Not only do we observe visual feature distinctions in this pre-cortical processing stage but there is also a laminar division of contralateral and ipsilateral visual field representations. An important feature of the LGN representation of the visual field is that a disproportionately large region of the thalamic space is dedicated to the foveal representation, a trait of visual processing which continues into the cortex. Initially considered a relay station in the retino-cortical pathway, the subdivision of the LGN as well as its connections with other cortical areas, indicates its role is more complex than traditionally thought. LGN output to visual cortex is in the form of four projection streams, containing segregated information of visual properties ready for cortical processing.

1.1.2 Cortical Processing

Once the visual signal reaches the cortex, a complex series of processing stages and interactions are undertaken in order to interpret and respond to the sensory input.

1.1.2.1 The visual hierarchy

At the cortical level, the visual system is arranged in a hierarchy (Young, 2000), denoted by its anatomical structure and reflected in its functional properties, with each level of the hierarchy differing in key aspects of processing.

Sensory input is projected from the LGN to primary visual cortex (PVC or also known as, V1). Anatomically, V1 lies at the occipital pole and occupies the banks of the calcarine sulcus. It is the first point of contact that visual signal has with the cortex, with virtually all visual tasks resulting in V1 activation (Grill-Spector & Malach, 2004a) and it therefore plays a key role in visual processing. As a result, it is one of the most extensively researched areas in visual cortex and indeed, cortex in general. Early work mainly involved invasive techniques on non-human mammals (Hubel & Wiesel, 1998), which, although not possible in humans and therefore could only be used to make inferences about human structure and function, did provide important and pioneering information about the visual system which still forms the fundamental knowledge basis that we work from today.

From V1, visual information transcends the visual hierarchy through higher visual areas, namely V2, V3, and so on; each processing different aspects of the information it receives (Livingstone & Hubel, 1987). V1 is considered a region in which the entirety of the visual signal is represented, albeit as a set of low-level features, before these properties are segregated and processed in parallel by the higher visual areas. This sequential stream of processing from lower to higher areas is considered feedforward or ‘bottom-up’ meaning sensory input moves up the hierarchy in a feedforward manner with increasingly abstract and holistic representations being formed at each level. This relatively straight-forward system formed the basis of traditional views of visual processing, however, as we will mention in more detail, it only forms part of what is now considered a dual processing account of the visual system.

1.1.2.2 Cortical organisation

Early investigation of the visual cortex, in particular V1, reveal distinct vertical organisation principles in the form of columns (Hubel & Wiesel, 1968;

Livingstone & Hubel, 1987). That is, neurons are organised into columnar structures, with all neurons within a column responding preferentially to similar stimulation features. Across both animal and human studies using a variety of techniques, cortical columns have been shown for numerous properties of visual processing such as ocular dominance and orientation selectivity (Hubel & Wiesel, 1974; Hubel, Wiesel, & Stryker, 1978).

Another vertical organisation feature of the cortex is that of layers. The cortex is a laminar structure, typically defined by six distinct layers, characterised anatomically by the abundance and type of neurons they contain. The superficial layers lie near the pial surface of the brain, and the deep layers extend towards the white matter. As well as differing in their structural features, the layers have also been shown to exhibit functional segregation, with different layers responding to different types of visual information as well as forming different functional connections within the layers themselves and to other cortical and sub-cortical areas. Advancements in MRI and fMRI now permit the exploration of the structure and function of laminar-resolution cortex greatly facilitating the research into layer-dependent functions.

The visual cortex also shows a high-level of horizontal organisation. The introduction of fMRI allowed larger regions of cortex to be investigated and reveals a series of maps throughout visual cortex (Benson et al., 2012; Brewer et al., 2005; Sereno et al., 1995; Wandell et al., 1994; Wandell & Winawer, 2011). In order to obtain such maps, subjects maintain central fixation while polar angle- and eccentricity-focused visual stimulation is presented to map the distance from the horizontal visual axis and the distance from the centre of the visual field, respectively. These maps represent the visual field topographically, meaning adjacent regions of visual space occupy neighbouring regions of cortex (via similarly retinotopically organised retinal and LGN representations). Multiple maps are found within different regions of visual cortex, each mapping a subsection or all of the contralateral hemifield, the boundaries of which can be determined through separation via horizontal and vertical meridians. Their representations of the centre of the visual field (i.e. the foveal representations) occupy a disproportionately large area of cortex, in a phenomenon known as cortical magnification (Duncan & Boynton, 2003; Qiu et al., 2006) and converge at the occipital pole. Retinotopic organisation is thought to be a property of earlier visual regions, however, higher-order visual areas have also been shown to be retinotopically organised (Grill-Spector et al., 1998; Hasson et al., 2003). For an in-depth look at cortical organisation, see (Rockland, 2017).

1.1.2.3 Receptive fields

Key components of the visual system are the receptive field properties of neurons. Classical receptive fields refer to the location and properties of the stimulus which preferentially causes a neuron to spike (Hubel & Wiesel, 1959). V1 neurons respond favourably to stimulation of a specific orientation in a specific region of the visual field. When such a stimulus meets these criteria, the neuron will fire. Therefore, the classical receptive field responds according to feedforward stimulation.

Firing rates of neurons however are not solely dependent on activation within the classical receptive field. Extra-classical receptive field properties mean that firing rates of V1 neurons are modulated by both lateral and feedback connections. The extra-classical receptive field is comprised of near and far surround regions which respond largely, although not exclusively, to input from lateral and top-down connections from higher visual areas (Angelucci & Bressloff, 2006). Response within the near and far surround tends to be suppressive in contrast to the excitatory-based classical RF but is not always. The presence of such extra-classical receptive fields mean that the spatial and temporal context in which a stimulus is presented can influence the receptive field properties of a given neuron (Wörgötter & Eysel, 2000).

Receptive field size increases along the visual hierarchy, with V1 neurons having relatively small receptive fields (Levitt & Lund, 2002) and higher visual areas having successively larger RFs. Thus, extra-classical RFs of lower visual areas are driven by stimulation of larger receptive fields from further up the hierarchy. As noted, cells in primary visual cortex are sensitive to orientation along with other low-level visual features (Goebel et al., 2012). This is conceptually logical as line orientation carries information about object boarders and edges, which is a salient feature of the visual scene. Other regions of visual cortex have receptive fields tuned to other visual features, giving rise to the functional specialisation within different areas of visual cortex.

1.1.2.4 Functional specialisation

The hierarchical model account of feedforward processing encompasses the processing of increasingly complex aspects of the visual scene starting with low-level visual properties and eventually creating a more holistic representation of the visual scene as the signal transcends the hierarchy.

Cortical processing begins with V1 which is traditionally considered a low-level feature detector (Petro, Vizioli, & Muckli, 2014), processing simple aspects of the visual input such as basic feature statistics. Areas outside of V1, known as extrastriate areas due to their anatomical distinction from striate PVC, are specialised in the processing of different aspects of the visual signal. V2 exhibits wavelength, orientation and direction selectivity within its neuronal population and is also thought to be somewhat involved in depth perception (Anzai & DeAngelis, 2010; Goebel et al., 2012). V3a and V3b/KO respond to contrast information and shape/moving contour information, respectively (Goebel et al., 2012). V4 also sometimes referred to as V8 (Hadjikhani et al., 1998) is considered a primary colour centre in the brain as it specialises in the processing of different wavelengths (Heywood & Cowey, 1987). V5/hMT+ shows high motion selectivity (Zeki, 2015), with preferential firing for moving elements within the visual array. Visual processing also extends beyond the occipital cortex to the parietal and temporal lobes. As such, these higher visual areas such as the LOC and PPA selectively respond to specific stimulus types such as objects and places (Cichy et al., 2011; O'Craven & Kanwisher, 2000).

It is important to note that these functional distinctions are not discrete. That is, although areas show preference for certain facets of visual information, they do not process these facets exclusively, with considerable functional overlap between regions. The system operates as a whole, processing every aspect of the visual scene in a sequential, yet parallel manner with a flow of information receiving functional contribution at every level.

Functional specialisation of the visual cortex extends to the presence of parallel processing streams; namely the ventral and dorsal streams(Milner & Goodale, 2008; Mishkin et al., 1983) (Goodale & Milner, 1992; Kravitz, Saleem, Baker, Ungerleider, & Mishkin, 2013; Milner & Goodale, 2008; Mishkin, Ungerleider, & Macko, 1983). These streams selectively encompass the specialist regions outlined above to process overarching properties broadly related to different aspects of visual processing. Work by Ungerleider, Mishkin and colleagues (Kravitz et al., 2013; Mishkin et al., 1983) defined two cortical pathways; the ventral stream which has an occipitotemporal trajectory and processes details related to object identification. They coined this the ‘what’ stream. The dorsal stream (also referred to as the ‘where’ stream by Mishkin and colleagues) is largely associated with object localisation and occupies regions spanning the occipitoparietal lobes. Later work by Milner and Goodale re-evaluated the function of these two streams and through patient-focused studies revealed the dorsal stream to be involved in visually guided action. Thus, they proposed the dorsal stream to be concerned more with ‘how’ to interact with objects rather than ‘where’ the objects are and subsequently termed the two streams as processing vision for perception and vision for action. Figure 1.1 depicts the main visual regions and where they lie in relation to each processing stream.

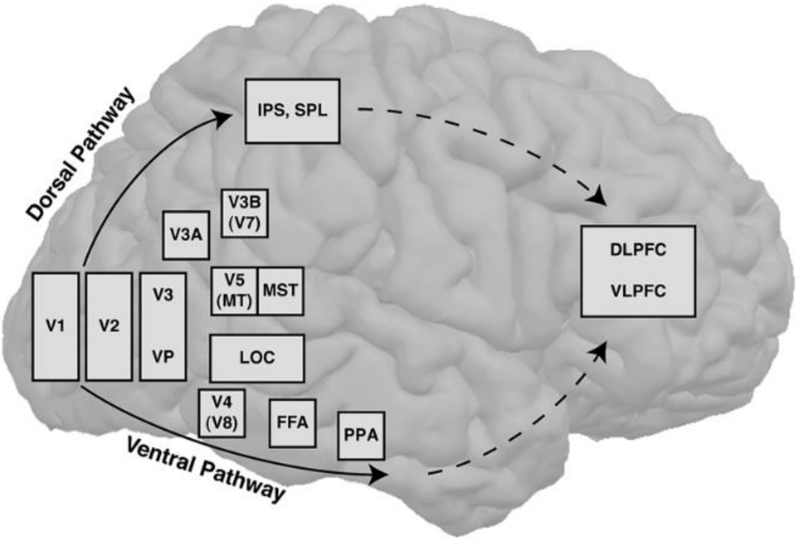


Figure 1.1 Key human visual areas. Prominent cortical areas of the human visual system. The solid lines indicate the dorsal and ventral visual pathways which encompass regions of the visual hierarchy. V denotes visual area; higher visual areas include MST – medial superior-temporal area; LOC – lateral occipital cortex; FFA – fusiform face area; PPA – parahippocampal place area; IPS – intrapartietal sulcus; SPL – superior parietal lobule; DLPFC – dorso-lateral prefrontal cortex; VLPFC – ventro-lateral prefrontal cortex. (Figure obtained with permission from (Goebel et al., 2012)**).**

1.2 Feedback in the visual system

1.2.1 What is meant by feedback?

In addition to the traditional sequential processing of feedforward external input, the visual system operates via another major stream which processes top-down information generated internally. The term top-down refers to aspects of cognition and behaviour that are not stimulus-driven (Engel et al., 2001). These two streams operate in parallel such that as sensory information is fed (forward) through the cortex via the retina and thalamus, top-down information from higher visual (and non-visual) areas is fed back down through the hierarchy in so that complex information in higher processing stages influences the processing at earlier stages (Gilbert & Sigman, 2007). In other words, the processing of sensory information is guided by influences from internally generated signals.

Evidence of feedback pathways can be taken from the architecture of the cortex itself. Regions within visual cortex are linked via reciprocal connections, highlighting a flow of information in both directions. This is evident throughout the cortex as a whole and even in sub-cortical processing within the thalamus (Gilbert & Li, 2013). Not only is the presence of this second information stream evident within the cortical architecture, its importance is also highlighted structurally with feedback connections outnumbering feedforward connections throughout the hierarchy (Friston, 2005). For example V1 contains more feedback than feedforward connections (Muckli & Petro, 2013). Similarly, most input to a given V1 neuron comes from different cortical regions, as illustrated in Figure 1.2. V1 receives more input from V2 than the LGN and upper layer pyramidal cells in V1 has approximately twice as many synapses with other cortical regions than V2 in macaques (Budd, 1998). Feedback connections from higher-order visual areas have also recently been mapped in cats (Pan et al., 2021). Cortical feedback projections have been found from the occipital face area which can modulate V1 activity during facial processing tasks (Petro et al., 2013). Topographic organisation of the intraparietal sulcus mirrors the retinotopic organisation of early visual cortex and the feedback projections from this area are thought to be implicated in attentional modulation (Greenberg et al., 2012). Additionally, feedback projections from auditory and multisensory cortices are thought to play a role in multisensory integration and spatial awareness (Clavagnier et al., 2004). Together these structural properties of the visual system highlight the importance of top-down influences on visual perception. It is also worth noting that while feedforward and feedback connections will be the focus of this discussion, a third neuronal input also operates during sensory processing in the form of lateral connections which connect neurons within a given cortical area. These have been shown to play a role in conveying spatial context information (Mayer et al., 2018) and integrating with feedback connections during image grouping and segmentation (Liang et al., 2017) Feedback connections on the other hand, connect neurons across cortical areas which are responsible for processing different aspects of visual information and their interaction with feedforward connections will form the focus of this thesis.

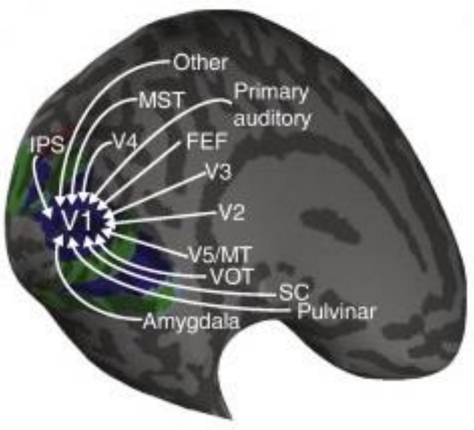


Figure 1.2 Areas connected to V1 via feedback pathways. Depiction of regions within and beyond the visual cortex which feed back to primary visual cortex. VOT – ventral occipitotemporal cortex; SC – superior colliculus; FEF – frontal eye fields; IP – intraparietal sulcus. Figure obtained and adapted with permission from (Muckli & Petro, 2013)**.**

This merging of the external and internal worlds during visual processing is now widely accepted as a template for how the cortex operates as a whole, and the brain is now commonly considered a parallel rather than serial processor (Singer, 2013), but the exact nature of how this dual-processing system is implemented and indeed the reasons why it is used are still being characterised. Below, I will attempt to address each of these questions using evidence from the current literature to piece together the properties and function of the feedback system.

Ultimately, we must understand what these feedback signals contain and how they are conveyed at the neuronal level in order to incorporate them effectively into models of how the visual system operates.

1.2.2 What does feedback do?

Generally, feedback influences the processing of incoming sensory input by conveying internally driven signals relating to the external information. Specifically, feedback encompasses a wide range of internally guided processes which modulate how the visual system responds to its visual environment in different ways. Processing everything in a feedforward manner would be computationally exhaustive. If the visual system operated in a purely bottom-up manner, every aspect of the visual scene (from individual objects and their relationship within and to the global structure of the scene) would have to be processed sequentially in a sensory-driven way. This would not only be energy- and time-consuming, but also does not seem efficient or even plausible in a rapidly changing environment that is the visual world. Indeed, it is advantageous to interpret the scene using guidance from previous experience than to continuously interpret a noisy signal (Panichello et al., 2013). It is therefore clear why a feedback system is needed, with the overall aim of top-down processing being to optimise information processing throughout the system (Teufel & Nanay, 2017) but the full extent of the benefits of such a system is yet to be clearly characterised.

A major role of top-down signalling is to provide predictions about the bottom-up information. In this sense, predictions refer to expectations about the current sensory environment formed through previous exposure and experience (Panichello et al., 2013). The terms prediction and expectation are used interchangeably throughout the literature with expectation describing a representation of what is predicted to occur and predictions describing a general orientation towards the future (Bubic et al., 2010). I will use the terms interchangeably throughout the thesis however it is worth noting that a consistent definition throughout the literature is not evident and a lack of systemisation in the terminology may influence the understanding of what exactly is being tested (Bubic et al., 2010). Predictive processing is based on regularities within our environment, (i.e. associations of objects and scenes that commonly co-occur), and how these predictive properties can influence perception (Bar, 2004). Its role is often considered a fundamental aspect of sensory processing and is implicated within many accounts of cortical function.

One way in which predictions can be formed is through contextual information. In the visual world, objects almost never appear in isolation; they are embedded within a visual scene. Surrounding information present within the visual scene serve as context for the objects within it and can greatly influence their interpretation. A scene is considered contextually coherent if it contains elements which commonly appear together with the right configuration (Bar, 2004). This coherence has shown to facilitate the perception of objects. For example, Palmer (Palmer, 1975) presented visual scenes followed by consistent or inconsistent objects. The objects were more easily identified when preceded by a contextually relevant scene. A few years later, using simple line drawings Biederman (Biederman et al., 1982), demonstrated that objects that are semantically consistent with their surrounding context are recognised more rapidly and accurately. Enhanced detection of consistent objects is also found by (Henderson & Hollingworth, 1999; Hollingworth & Henderson, 1998) This effect was extended to natural viewing conditions by Mack & Eckstein (Mack & Eckstein, 2011), who found visual search to be more efficient for expected versus unexpected stimuli.

Contextual information has been shown to facilitate the recognition of other objects which share the same context (Bar & Ullman, 1996) and recent work by Bar and colleagues (Afiki & Bar, 2020) has shown that associated (versus un-associated) pairs of images resulted in improved performance in several visual perception-related tasks. The human trait of seeking coherence within our visual environment and evaluating probabilities of object-scene co-occurrence is emphasised by Sauvé and colleagues (Sauvé et al., 2017) in a study which looked at both behavioural and electrophysiological correlates of contextual effects, based on previous work by Bar and colleagues (Bar, 2004). As well as a behavioural advantage of object-scene pairs which had a high probability of co-occurrence, they also found differences in event-related potentials following stimuli with different levels of likelihood of co-occurrence between the object and scene. This highlights both a behavioural and neurological effect of contextual modulation.

Contextual modulation of objects highlights a key function of top-down predictive processing. However, in order for this process to seem more ecologically valid, it must also extend to the perception of scenes in general rather than simply the objects within them (Bar, 2004; Greene, 2013). We view the world as a dynamic flow of visual arrays, rich in visual information and therefore scenes are more representative of our visual experience. In addition to scene information facilitating object processing, object recognition has been shown to aid scene perception (Davenport & Potter, 2004). Caddigan and colleagues (Caddigan et al., 2017) demonstrated that ‘good’ scenes (those which are more representative of their category) were more easily detected than ‘bad’ scenes (those less representative of their category) even when the differences were not task relevant. Thus, contextual modulation is implicated in both object and scene perception, although arguably these processes are not mutually exclusive within visual processing itself and therefore should reflect an overall influence of visual recognition by context.

Information in the context is also useful when elements of the visual scene are ambiguous. For example, (Bar & Ullman, 1996) presented portions of a visual stimulus which could contain ambiguous or unambiguous features. They found that the presence of key identifiable features aided the identification of the ambiguous elements. An ambiguous object could be interpreted in different ways depending on the contextual information it is presented with. For example, a blurred object could be recognised as a hairdryer if portrayed in a bathroom scene or as a drill if presented within the context of a garage or workshop (Bar, 2004). This effect of context on ambiguity has been replicated in computer vision studies (Oliva & Torralba, 2003; Torralba, 2003).

Effects of predictive processing in relation to ambiguity are also evident in perceptual illusions which prompt perceptual uncertainty or offer different ways to interpret the visual input. A classic example of this is binocular rivalry in which two different images are presented to each eye. Instead of merging these two percepts, each takes it in turn to dominate perception. Contextual information has been shown to influence the experience of binocular rivalry by modulating the dominating percept (Denison et al., 2011; Sobel & Blake, 2002).

The effects of contextual modulation are bi-directional. Contextual information that is incoherent with the objects presented, can disrupt or slow down their recognition (Biederman et al., 1982). Predictions derived from contextual information can also 'taint' the perception of visual stimuli (Bar, 2004), affecting the processing of sensory information but not necessarily in a perceptually relevant or beneficial way. This is evident through phenomenon such as false memories (Miller & Gazzaniga, 1998), boundary extension (Intraub et al., 1996) and change blindness (White et al., 2016).

Evidence of top-down modulation within the visual cortex can also be found at the neuronal level. Predictable information has been shown to enhance activity in the visual cortex (e.g. Mannion, Kersten, & Olman, 2015). Conversely, enhanced activity has also been observed for unexpected stimuli (Alink, Schwiedrzik, Kohler, Singer, & Muckli, 2010; Joo, Boynton, & Murray, 2012; Kok, Van Lieshout, & De Lange, 2016), meaning we are yet to reach a consensus over how predictions influence cortical activity at the univariate level. Expectations have been shown to bias representations in visual cortex (Kok et al., 2013), with stimulus templates reflected in the pattern of activity in PVC being present when the stimulus is expected (Kok, Failing, & de Lange, 2014).

Cortical representations have also been shown to be enhanced by the surrounding context for individual written letters (Heilbron et al., 2020), a neural feature which presumably facilitates the ability to read. A link between the neural activation and resulting percept was also found by (Joo et al., 2012), who found that subjects were more likely to see an object that deviated from the surrounding context in conjunction with greater cortical activation for such stimulation. Rideaux & Welchman (Rideaux & Welchman, 2019) were able to decode information about perceived depth from the neural representation within V1, again linking neural representation with the resulting percept. Similarly, it has been found that objects perceived as different sizes occupy different amounts of primary visual cortex in their neural representations (Murray et al., 2006). They demonstrated using fMRI that objects that were perceived to be distant (and therefore perceived as occupying a larger portion of the visual field) activated a larger portion of the visual cortex relative to objects (of the same angular size) perceived as closer (and subsequently perceived to occupy a smaller portion of the visual field). This highlights a contextual effect, in the form of contextual depth cues, in early visual cortex.

Contextual information has been shown to bias perception in the absence of direct visual input, implying that top-down influences can exert an effect without any bottom-up stimulation. Haynes & Rees (Haynes & Rees, 2005) were able to decode the orientation of invisible stimuli from the response pattern in visual cortex. In a similar vein, Smith & Muckli (Smith & Muckli, 2010) were able to decode information about the surrounding context in an unstimulated area of V1. They achieved this through occluding a portion of the visual scene, effectively eliminating feedforward input to the corresponding region of visual cortex. Several other studies have also employed occlusion paradigms and found evidence of top-down signalling in the absence of bottom-up input (Ban et al., 2013; Chen et al., 2018; Muckli et al., 2015; Sugita, 1999). These findings imply feedback signals can carry information about the surrounding context, a notion supported by Williams and colleagues (Williams et al., 2008) who found foveal cortex to contain information about peripheral stimulation. Further evidence of cortical activation from indirect stimulation this comes from the attentional blink paradigm (Lindh et al., 2019) studies looking at the cortical correlate of the blind spot (Chen et al., 2017), illusion studies (Chemla et al., 2019; Dekel & Sagi, 2020; Muckli et al., 2005; Vetter et al., 2012), as well as mental imagery studies (Bergmann, 2019; Gosselin & Schyns, 2003; Keogh et al., 2020; Smith et al., 2012).

The majority of the findings discussed implicate V1 and early visual areas within predictive processing and contextual modulation. These processes have recently been shown to be driven by low-level features (Lauer et al., 2018), (a known driver of activity in PVC) but modulated by scene complexity (Groen et al., 2018). Evidence of these processes are also found in higher visual areas (Faivre et al., 2019), suggesting feedback is influential on many levels of the feedforward hierarchy. The feedback signals themselves are thought to originate from higher visual areas but have also been shown to implicate other sensory areas such as auditory cortex (Meijer, Montijn, Pennartz, & Lansink, 2017; Petro, Paton, & Muckli, 2017), highlighting a potential role within multisensory integration.

Top-down predictions are also considered to play a role in other higher-level processing such as working memory and attention (Mehrpour et al., 2020). Activation of working memory is required to maintain predictions which need to be updated and held within WM as the sensory information is further processed. Knowledge may be stored as early as V1 and then reactivated by top-down feedback signals (Petro et al., 2014). Pratte & Tong (Pratte & Tong, 2014) demonstrated working memory representations being stored in V1 through the ability to decode stimulus information during a retention period.

The relationship between top-down signalling and attention is bi-directional, with attention being attracted to salient stimuli (so-called bottom-up attention), as well as being voluntarily directed by top-down control (Connor et al., 2004), particularly by prefrontal and frontal areas (Katsuki & Constantinidis, 2014; Paneri & Gregoriou, 2017). Gilbert & Li (Gilbert & Li, 2013) outline evidence for different types of attention influenced by top-down control. Firstly, spatial attention, most commonly associated with a gain control mechanism whereby relevant or salient stimuli are captured by attention at particular points in the visual field or conversely, irrelevant stimuli are suppressed (Motter, 1993; Posner et al., 1980). In contrast to a searchlight-based control of attention, attention can also highlight object properties or features of the visual field with common properties. Therefore, attention appears to be a complex and multi-level mechanism of top-down control which ultimately influences how the visual field is processed. For more detailed discussions see (Muckli & Petro, 2013) and (Gilbert & Sigman, 2007).

Attention could be used by top-down predictions exert an effect across processes involved in visual perception. For example, top-down effects could prime the system to attend to specific features (Serences & Boynton, 2007) or areas in space but in turn, expected or unexpected stimulation could capture attention. Endogenous attention in particular is referred to as top-down, goal-driven attention (Maclean et al., 2009) operating via the ventral pathway using predictions and expectations to guide attention in contrast to exogenous attention which responds to external factors such as stimulus properties. It is therefore difficult to disentangle the processes of expectation and attention as there is considerable overlap. Attention can mediate contextual influences and thus in turn influence predictions. Although often considered interchangeable, expectation and attention should be kept apart according to Panichello and colleagues (Panichello et al., 2013), an opinion based on work by Summerfield and Egner (Summerfield & Egner, 2009) who define the distinction between attention and expectations. Attention prioritises processing on the basis of relevance whereas expectation restrains interpretation depending on likelihood. In other words, attention may determine the response to the stimuli but predictions may provide an anchor to modulate this response (Hsu et al., 2014). Summerfield and Egner detail differences between the two processes but also outline how they overlap and interact. An ERP study by Marzecova and colleagues (Marzecová et al., 2017) show differential effects of both predictability and attention but also how these factors integrate in visual processing. Through a meta-analysis Ficco and colleagues (Ficco et al., 2021) found overlap between the predictive coding and attentional networks. These works highlight both the separation and integration of these two important mechanisms in visual processing. Although the picture of how these processes interact is not yet clear, it seems that the endogenous attention system acts as a channel for predictive processing to operate. Exogenous attention is then used to provide the bottom-up input. An example of which can be found in a study by Macaluso & Doricchi (Macaluso & Doricchi, 2013).

Overall, the role of feedback in the visual system is complex with many attributes that need taking into consideration before we can build an effective model of its function. We know it is implicated in processes such as object and scene perception, working memory, attention, predictive processing, contextual modulation and multisensory integration, as well as many more within and beyond the visual system. Many features of these processes overlap and could reflect an overarching influence of feedback or the contribution of many sub-processes and the interplay between them. What we do know is that feedback plays a crucial role in cortical function, and we are potentially just scratching the surface of discovering the extent of its nature.

1.2.3 How is feedback implemented in visual processing?

We have discussed the varying processes influenced by top-down processing and can appreciate the importance of the feedback system, but how are top-down signals implicated in visual processing and how do they interact with the counter-current stream of feedforward input? Addressing these questions taps into an area of research that is rich in theoretical concepts encompassing a wide range of hypotheses and possible mechanisms underlying cortical function. I will highlight a few of the key ideas that attempt to explain the way in which feedback signalling is implemented throughout the visual system. It is important to note that these concepts are not mutually exclusive, some encompass or work in conjunction with others. While there is plenty of evidence supporting each notion, a clear consensus on how feedback operates within the cortex is yet to be established.

Attempts to ascertain how feedback is employed within the brain start with some more abstract conceptualisations. For example, Bar (Bar, 2004) suggests that contextual modulation occurs through context frames, which are structures of contextually consistent information created from experience. These context frames are then activated by cues from the feedforward information and trigger the top-down facilitation of object recognition. Bar (Bar, 2007) goes on to develop a three-part model of contextual modulation in which information from the sensory input develops analogies which are linked to memory. These analogies initiate associations which are used to create predictions about the information being presented. Bar also proposes that context may activate top-down predictions which are used to lower the perceptual threshold, in turn increasing sensitivity to relevant information. Increased sensitivity is also outlined by Serences & Boynton (Serences & Boynton, 2007) who suggest that feature-based attention, guided by top-down processes, increases sensitivity to behaviourally relevant information.

Another way in which top-down processing is activated is thought to be through the type of information that is extracted from the bottom-up input. Specifically, it is proposed that the visual system rapidly extracts a coarse representation of the visual scene (Bar, 2003) which is used to activate context frames used for further processing (Bar et al., 2006). This coarse-level extraction comes in the form of low-spatial frequency information which conveys enough information about global scene properties to activate high-level predictions. These perhaps originate as high as the orbito-frontal cortex (OFC), which shows early activation, reminiscent of expectations. The quick extraction of LSF information provides the gist of the visual scene (Schyns & Oliva, 1994) and activates schemas stored in memory, which is followed by slower processing of the HSF information conveying fine-grained information to fill in details. This is similar to Bar's proposal (Bar, 2007) whereby LSF information activates analogies stored in memory however Bar expands this to explain how analogies may then be used to eventually form predictions about the visual input.

Other theories encompassing neuronal function-based hypotheses posit several processes involved in visual perception as potential mechanisms for how feedback exerts an effect. These include surround suppression (Er et al., 2020) in which a neuron's response is reduced when stimulated from outside its classic RF, or mechanisms that operate on a much larger scale such as synchronised neural oscillations at distinct frequencies for feedforward and feedback information (Bastos et al., 2015) or a general network state mechanism (Zagha et al., 2013). Smaller scale notions propose that rather than point processors, individual neurons themselves instead act as adaptive processors, dynamically combining top-down and bottom-up input (Gilbert & Sigman, 2007). A more simplistic neuronal account of top-down modulation points to receptive field characteristics of V1 neurons. Indeed, Gilbert & Li (Gilbert & Li, 2013) stipulate that understanding how cognitive influences affect neuronal function requires an understanding of the receptive field. So, when addressing how top-down influences exert an effect, one should consider that this may be through RF properties that allow stimuli outside of the classical receptive field (i.e., stimuli which is unable to elicit a response alone) to drive neural activity when combined with specific information inside the classic RF. In other words, information from large parts of the visual field can be integrated within a single RF response profile due to top-down modulatory effects.

Beyond the computational capabilities of individual neurons, neural circuits have been implicated in the integration of feedback and feedforward information, such as canonical microcircuits within cortical columns (Bastos et al., 2012;

Keller & Mrsic-Flogel, 2018; Wibral et al., 2017) or circuits connecting cortico- and either thalamic (Mumford, 1991) or (higher) cortical (Mumford & Mumford, 1992) inputs. Furthermore, accounts of cortical function as a whole have been used to explain the process of top-down processing. Hierarchical Bayesian inferencing is a popular framework encompassing several relevant theories of visual function (Lee & Mumford, 2003, 2007) but it is important to note that its central notion applies to general brain function as is not limited to that of the visual system. The core understanding of Bayesian inference posits that loops combining feedforward and feedback information apply probabilistic inference to merge the available sensory information with experience-driven contextual priors (Lee & Mumford, 2007). Feedforward input triggers hypotheses about the sensory input whereas feedback information provides predictions to generate the hypotheses. Each pathway is used to systematically update and modify the signal; thus, the inference of the visual input is determined by information from both. Lee & Mumford outline a potential problem with this notion in that if an incorrect interpretation is implemented in either pathway, the other will be updated accordingly which may lead to an invalid conclusion being drawn. Therefore, they suggest it may be beneficial to maintain several hypotheses in parallel to circumvent this potential issue.

A theory formed on the basis of Bayesian inferencing is predictive coding (Hohwy et al., 2008; Kwisthout et al., 2017; Shipp, 2016; Spratling, 2010). Predictive coding integrates the role of feedback and feedforward signalling by proposing that feedback signals provide predictions about the feedforward input gleaned from statistical regularities encountered through previous experience. These predictions are then compared to the input from the feedforward signal and any disparities (termed prediction errors) are fed to the next level of the hierarchy (Rao & Ballard, 1999). This comparison of predictions and input (referred to as a matching inhibition mechanism by (Clark, 2013)) occurs at every processing stage and the errors are continuously reduced with every hierarchical step. Information which is compatible between the two streams is ‘explained away’ by cancellation (Rao & Ballard, 1999) or attenuation (Alink et al., 2010; Fang et al., 2008) of the signal, so that only information which required further processing advances to the next stage. This is a concept underpinned by the principle of minimising free energy (Friston, 2005; Friston et al., 2015; Gershman, 2019).

In contrast to predictive coding accounts of cortical function, which ultimately imply that signal is enhanced only when feedforward information does not match the top-down predictions, other theories propose the opposite. Theories such as adaptive resonance theory (Grossberg, 2013) and coherent infomax (Kay & Phillips, 2011) suggest that neural activity is enhanced when top-down and bottom-up signals are compatible. Once such neural account of this enhancement is that of apical amplification (Kay & Phillips, 2018; Phillips, 2017; Phillips et al., 2016). Apical amplification stipulates that feedback and feedforward integration can occur at the level of individual neurons. In particular, layer 5 pyramidal cells are implicated in this function whereby cellular response to feedforward stimulation is amplified if the top-down signal is consistent. A mechanism in which this may be implemented is via BAC firing, whereby coincidental input to the apical dendrites (via feedback signalling) and the cell soma (via feedforward input) triggers a cascade of action potentials via calcium spiking channels, a process shown to occur within rodent pyramidal cells (Larkum, 2013; Major et al., 2013). BAC propagation is therefore a strong candidate mechanism for apical amplification and other theories which propose an increase in neural response to consistent information. Figure 1.3 and Figure 1.4 illustrate how feedforward and feedback signals are integrated at the neuronal level.

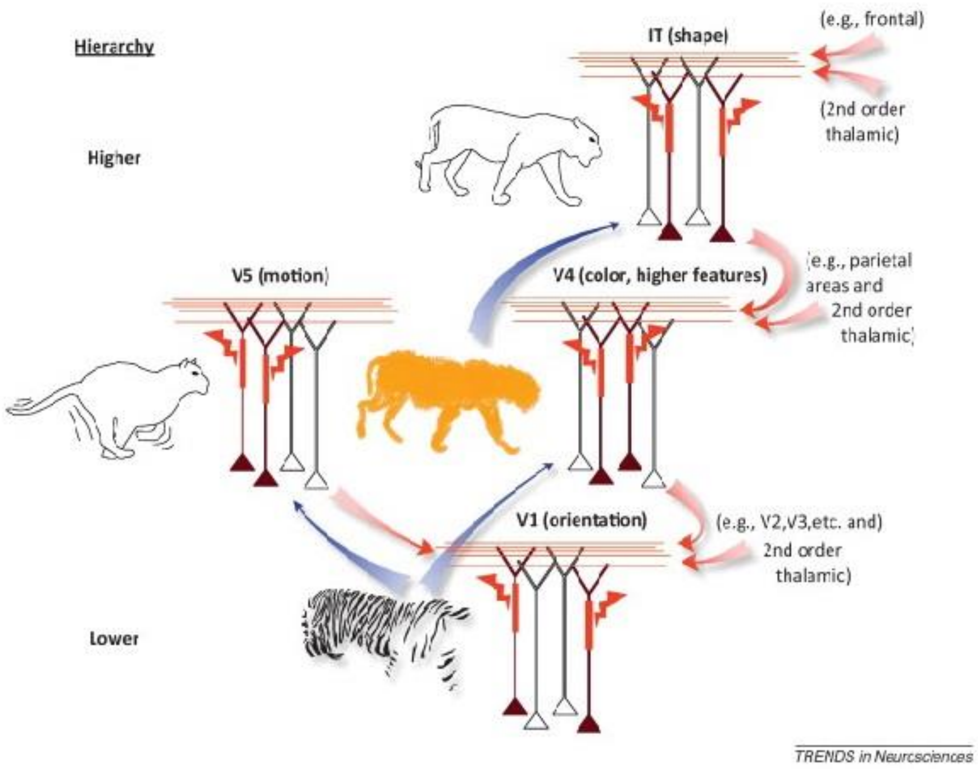


Figure 1.3 A cellular mechanism for integrating feedforward and feedback signals in object recognition. Low-level features are encoded in primary visual cortex and transcend the hierarchy. Feedback inputs from various sources (red arrows) are carried by horizontal fibres that synapse on the distal tuft dendrites. Feedback therefore informs lower areas about higher-level representations to serve processing. BAC firing serves as a possible mechanism in which top-down information influences bottom-up processing. This schematic shows a simplified representation of the visual system, highlighting key areas and features. Figure obtained with permission from (M. Larkum, 2013)**.**

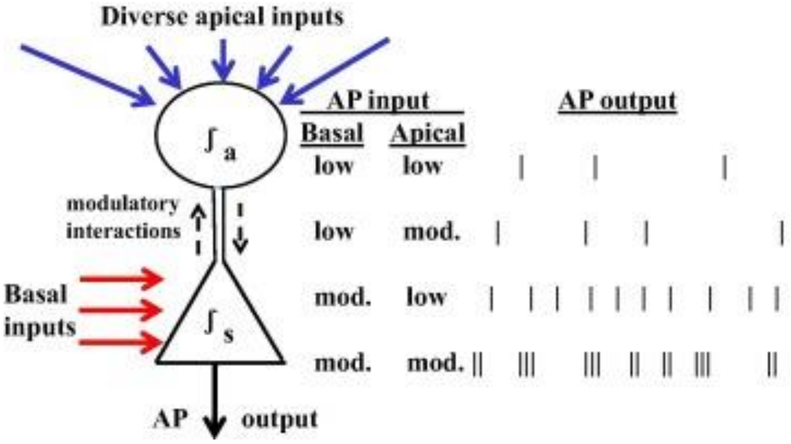


Figure 1.4 Integration sites of pyramidal cell neurons. The two integration sites of pyramidal cells $\int s$ shows the somatic integration site which generates action potentials. $\int a$ is the apical integration site which when combined with sufficient activation at the somatic integration site can convert a single action potential into multiple potentials. The potential outputs are depicted on the right-hand side. Figure obtained with permission from (Phillips, 2017)**.**

Rather than amplifying or suppressing coherent feedback and feedforward signals, a third possibility is that consistent predictions are used to sharpen the representation of upcoming stimuli. That is, that consistency between expectations and sensory input has been shown to reduce response amplitude, while improving the stimulus representation within the BOLD response pattern in early visual cortex (Kok et al., 2012). This sharpening of representation has recently been extended to deep-learning applications (Abdelhack & Kamitani, 2018).

1.3 Thesis rationale

With the many advantages of top-down processing highlighted, and a particular emphasis being found on contextual modulation, this thesis broadly aims to investigate how top-down and bottom-up processing streams integrate within the visual system to influence perception. More specifically, we aim to explore how top-down contextual information can influence the processing of degraded feedforward input.

To do this, we will harness the capabilities of occluding a portion of the visual scene in being able to isolate feedback signals. Investigating the role of feedback in general is challenging due to its implicit integration with feedforward information. Occlusion, however, permits the separation of feedback from feedforward information in order to test their individual contributions to visual processing. The occlusion format we will use will follow that of Smith & Muckli (Smith & Muckli, 2010), in which feedback signals carrying information about surrounding context were found in unstimulated primary visual cortex, suggesting the visual system compensates for a lack of input using top-down predictions. To develop these exciting findings further, we aim to test how these feedback signals may be used when degraded feedforward information is present. In a real-world environment this could be akin to situations in which perceptual ambiguity is present and contextual information may be used to interpret the signal, for example, like looking through a frosted pane of glass – feedforward input is weak and therefore contextual information must be used to ultimately determine its meaning. By including weak feedforward information within our occluded region, we reintroduce the amalgamation of the two signals to an extent, but in doing so we hope to gain an insight into how the top-down signals can help us interpret the weak, feedforward ones. Smith & Muckli's findings have shown that feedback signals are present and now we want to explore how these signals may be used during visual processing. It is important to note that other salient processes involved in top-down control will naturally play a role in our experiments, however we are not specifically searching for the influence of attention and/or working memory here so manipulation of these mechanisms does not form part of the experiments outlined. However, their implications will be discussed in brief where relevant.

We aim to address this thesis aim on two levels: psychophysically and neuronally. We firstly want to test whether manipulation of the contextual information used to generate the feedback signals can result in different behavioural outcomes. That is, do the expectations derived from the feedback signals, and consequently the predictability of the feedforward information result in different behavioural responses. Based on findings within the literature, we expect to see a behavioural advantage of contextually coherent information, and perhaps a disadvantage of incoherent information.

We will then investigate this effect by looking at the cortical consequences of such manipulation. Do relevant predictions alter the neuronal response of the visual cortex relative to irrelevant predictions? Previous findings have revealed mixed neuronal responses to contextually coherent or incoherent information which is reflected in theories of cortical function which encompass neural mechanisms which could either enhance or suppress signal for expected or unexpected stimuli. We anticipate neural differences between the two types of signals but the direction of which less clear.

More specifically we will aim to address the following questions: In Chapter 2 we will investigate how top-down and bottom-up inputs interact to influence the processing of degraded feedforward information at the behavioural level. In Chapter 3 we will test the effects of consistency at the neuronal level by exploring whether top-down predicted information can amplify degraded feedforward input within primary visual cortex. Furthermore, Chapter 4 will outline a project proposal which aims to expand these behavioural and neuronal findings to explore a contextual effect on degraded information at a higher resolution, enabling a greater insight into the neuronal underpinnings of the integration of feedback and feedforward inputs.

By incorporating feedforward input into a traditionally top-down driven paradigm, we hope to elucidate the role of each of these information streams within visual processing and hope to add to the growing body of evidence highlighting the importance of feedback signals within cortical function.

2 Chapter 2

2.1 Abstract

The recurrent processing of the visual system is well recognised but still yet to be fully understood. The way in which top-down and bottom-up inputs are integrated in the processing of the visual scene is an important aspect of visual processing that requires thorough exploration. Here, we present a series of psychophysical experiments which aim to investigate how top-down and bottom-up information is integrated to process degraded feedforward information. We employ a partial occlusion paradigm in which we present natural visual scenes with the bottom right corner shown at low contrast so that the information that it contains is degraded compared to the full contrast surrounding information. This surrounding information either matches or does not match the low-contrast region meaning the contextual information provided is either relevant or irrelevant. Through our experiments we then explore how well subjects are able to detect and/or identify the low-contrast scene.

Results reveal a consistency advantage in terms of recognition only, in that subjects are better able to identify the content of the low-contrast information when it is surrounded by consistent contextual information. When the context is inconsistent, performance decreases. This finding is revealed through both accuracy-based and Signal Detection Theory-based analyses which account for any response bias that could be driving the effect. While a response bias toward the surrounding context appears to contribute to the consistency effect we observe, it does not account for it wholly and thus our results provide evidence of a contextual modulation upon degraded feedforward information. This gives insight into the way in which feedback and feedforward information is used at the behavioural level within visual processing when input is degraded or ambiguous.

2.2 Introduction

2.2.1 Theoretical Background

The visual world can present an unlimited number of individual scenes, containing an infinite array of objects differing in numerosity and spatial configuration. Our ability to recognise thousands of object categories within scenes despite these variations is a capability of the human visual system still unmatched by computer algorithms (Oliva & Torralba, 2007). Processing every visual scene in a bottom-up, sequential manner would use an unreasonable amount of neuronal processing, however, the visual system operates using a two-way processing system whereby bottom-up input is processed with the help of top-down information from other (visual and non-visual), (Felleman & Van Essen, 1991; Markov et al., 2014; Vezoli et al., 2021).

A crucial role of this top-down flow of information during visual processing is to provide predictions about the visual environment (Petro et al., 2014). Predictions are formed from our past and current experience of the world and serve to help us efficiently process current and upcoming visual input. One way in which these predictions can be formed is by using contextual information available within the visual scene (Bar, 2004).

There is a wealth of literature which illustrates how contextual information can facilitate the processing of objects and scenes. Typically, contextual information facilitates the processing of objects which are consistent with the scene they are presented in. Contextual information has also been shown to facilitate the processing of ambiguous objects. (Bar & Ullman, 1996), found that ambiguous objects were more easily identified when surrounded by reliable contextual information. Additionally, inconsistent context has been shown to inhibit effective processing (Biederman et al., 1982; Green & Hummel, 2006; Palmer, 1975). Other experiments, however, have even found a facilitatory effect of inconsistent context in object processing (Brockmole & Henderson, 2008; Henderson & Hollingworth, 1999).

One way to assess the presence and/or content of these feedback signals is by occlusion which provides the motivation behind the paradigm used in this project. Smith & Muckli (Smith & Muckli, 2010) adopted a partial occlusion paradigm (see Figure 2.1) to isolate feedback signals to a region of early visual cortex. They presented a series of visual scenes and occluded the bottom right corner, eliminating feedforward information in this region of the visual field. During either a blocked design or a rapid event-related design, subjects were asked to maintain fixation and perform either a colour change detection task or a one-back repetition detection task, respectively. In the corresponding (unstimulated) region of early visual cortex (both V1 and V2), using a multivariate pattern classification-based analysis they found signals containing information about the surrounding scene; evident through the classifier's ability to accurately decode the surrounding scene within this region of cortex. This effect persists across both experiments, in both areas (V1 and V2) and using different types of classifiers and classification techniques. In a further study in 2015, Muckli and colleagues (Muckli et al., 2015) extended the exploration of this feedback effect to a layer-specific analysis. Using ultra-high field fMRI, they again found evidence of feedback signals in occluded visual cortex.

These results therefore robustly imply that contextual information is fed back to cortical areas deprived of feedforward information and used even in the absence of bottom-up input. We want to extend this finding by investigating how context influences processing when some degraded feedforward information is present. Although we plan to investigate this neuronally, we initially want to explore whether we can find a behavioural influence of feedback information on the processing of degraded input. Therefore, unlike these previous occlusion studies, our task will explicitly test for a contextual effect of surrounding ('feedback') information.

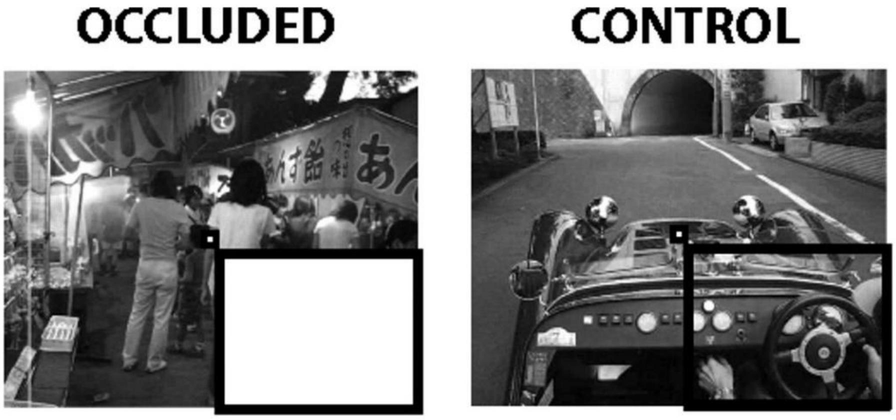


Figure 2.1 Partial occlusion paradigm used by Smith & Muckli. The two conditions used by Smith & Muckli (2010) (adapted with permission). Left – occluded: the bottom right quadrant of the image is replaced by a uniform white field. Right – control: the bottom right quadrant contains the same scene as the remaining image. Black lines are for illustration purposes to highlight the regions.

The main question this chapter will address is: how do top-down and bottom-up inputs interact to influence the processing of degraded feedforward information? We will investigate this through a series of psychophysical experiments which will explore different aspects relating to this central aim.

2.2.2 What we mean by top-down inputs

Top-down signalling refers to the information fed back to lower visual areas to aid the processing of upcoming sensory information, or, more generally, cognitive influences on earlier steps in processing (Gilbert & Li, 2013). As reviewed above, feedback may contain contextual information extrapolated from the scene and combined with prior knowledge to make sense of the feedforward information currently available. In our experiment series, we will manipulate the reliability of these predictive signals by altering their contextual information. As in the partial occlusion paradigm by Smith & Muckli, outlined above, the main portion of our stimulus images will contain full-contrast contextual information. In our case, the occluded region will also contain degraded feedforward information. This will allow us to compare feedforward and feedback influences by manipulating the consistency between the scenes depicted in the surround and occluded regions, in turn manipulating the reliability of the context and thus the predictability of the degraded information. In other words, we will assess subjects’ ability to interpret the content of the occluded region when surrounded by either consistent or inconsistent context.

A further effect we wish to test is whether contextual facilitation can extend beyond simple image features such as line continuation within our paradigm. Line continuation plays a crucial role in facets of perceptual processing such as boundary extension (Gottesman & Intraub, 1999) and amodal completion (Murray et al., 2004; Nanay, 2018). Recent work by Morgan and colleagues (Morgan et al., 2019) has shown that predictions reliably follow contours and boundaries using information available in the context. Here, the consistency between the surround and occluded information will be determined on different levels of categorisation, meaning both portions of the stimulus image may not be parts of the same image but instead may be considered consistent at the subordinate or basic category level. Basic level categorisation is the level at which the most relevant conceptualisation of a category is found. For example, *'dogs'* would be classed as a basic-level category. Subordinate categorisation further divides the basic categories. Using our example, *'Spaniel'* would be a subordinate level category. Thus, our stimulus images will be consistent along this continuum of categorisation.

Another important feature in object and scene recognition literature is spatial frequency information. Findings from (Bar et al., 2006) suggest low spatial frequency information is sufficient to drive an early predictive response about the scene. This notion is supported by early work from Schyns & Oliva (Schyns & Oliva, 1994) who suggest visual processing occurs in a coarse to fine manner, of which low spatial frequencies could contribute to the coarse information referred to in their findings. Bar then proposes that higher spatial frequencies drive a later response to confirm or rule out predictions based on the gist information gleaned from the lower spatial frequencies processed earlier. The importance of low spatial frequency information in early scene processing and a distinction between different spatial frequencies in visual processing generally is supported by more recent work (Dima et al., 2018; Kauffmann et al., 2014; Kihara & Takeda, 2010; Mu & Li, 2013). We will therefore assess the consequence of manipulating spatial frequency information in both the context and occluded region in turn, within our paradigm. We expect that removing low-spatial frequency information in the context will reduce the ability to infer the identity of the degraded input. How this interacts with the contextual influence will help us determine what type of information is portrayed in the feedback signals used to facilitate performance. Limiting spatial frequencies in the occluded region will increase task difficulty so may enhance or eliminate any contextual modulation we observe.

2.2.3 What we mean by Bottom-Up inputs

The bottom-up aspect of our central research question pertains to the degraded feedforward information in the occluded region. Previous occlusion studies have occluded one portion of the visual field entirely, eliminating feedforward input to a region of visual cortex. Our occluded region, however, will contain degraded scene information. How does the level of degradation of the feedforward input modulate the contextual effects?

We will degrade the scene information in the occluded region by lowering its contrast. Reducing the contrast of an image removes fine grained information however objects can still be identified with degraded visual conditions if global information is preserved (de Cesarei & Loftus, 2011). At the lowest levels of contrast, feedforward information may become invisible.

There is a multitude of research showing how top-down information can influence the processing of perceptually invisible stimuli. For example, visual illusions such as the Kanzsia triangle (Kanzsia, 1979, found in (Wang et al., 2012)) and motion induced illusions such as apparent motion (Muckli et al., 2005) arise from top-down predictions which create the illusion of either contours or a motion trajectory, meaning some of the information processed has no direct retinal input. This finding is also replicated in instances of perceptual rivalry with the filling in of information in the absence of retinal input (Chen et al., 2017). Context also influences illusory percept when the stimulus does stimulate the retina, for example in binocular rivalry, information in the context influences which image is selected (Denison et al., 2011) or maintained (Sobel & Blake, 2002) in conscious perception. Other studies have explored contextual modulation when the context itself is rendered invisible, although some effects have been found, they are less pronounced than if the stimulus itself is beyond conscious perception (Biderman et al., 2020; Harris et al., 2011).

At the neuronal level, the orientation of invisible stimuli can still be predicted from activation patterns using fMRI (Haynes & Rees, 2005) and global context has been shown to influence local processing as early as V1 (Rideaux & Welchman, 2019; Smith & Muckli, 2010). Faivre and colleagues (Faivre et al., 2019) found different neural signatures of stimuli that were congruent versus incongruent with their context when the stimuli were visible but this did not persist when the stimuli were invisible, indicating a disparity between information that is or isn’t consciously perceived. By testing perceptual threshold, we can investigate the role conscious perception plays in contextual modulation at the behavioural level.

We will firstly include an array of contrast levels with a wide range spanning from very low to a level at which the scene is fairly visible but is still substantially degraded compared to the context. This will enable us to gauge the behavioural consequence of the processing of varying degrees of reduced feedforward input and how the predictions based on high contrast information influence the ability to process low-contrast information.

We will then limit contrasts to around the visual threshold and test the influence of top-down predictions on the processing of this threshold-level information. This will allow us to investigate whether consistency of contextual information influences the detection and/or identification of scene information that is not necessarily perceptually visible.

If top-down predictions enhance or amplify consistent information, then we would expect a beneficiary effect even around threshold level. Evidence supporting the notion of amplification comes from neuronal architecture and function. Apical amplification is a mechanism proposed to serve this function in which top-down information can enhance the processing of feedforward input by increasing firing rates of neurons receiving both types of input (Phillips, 2017). Although we know that top-down information influences how the brain interprets bottom-up input and there are proposed mechanisms to explain how this occurs, we are still far from knowing exactly how this process manifests cortically and behaviourally and this series of experiments aims to contribute to the understanding of the latter.

2.2.4 What we mean by processing

As we focus on how top-down information influences the processing of bottom-up input, it is important to clarify what aspects of visual processing we refer to and how we plan to investigate this.

We will firstly consider two important aspects of visual processing: detection and recognition. Detection refers to the ability to determine whether or not a signal is present. Recognition refers to the ability to accurately determine what that signal is. Although they can be considered relatively separate processes (Straube & Fahle, 2011), they are considered interlinked (Swets et al., 1978).

These aspects of processing will be tested separately and in combination throughout our experiment series, subjects will be tasked with both a detection and identification task to investigate whether contextual information exerts a similar effect on each.

Additionally, we will focus not only on performance accuracy but we will also employ Signal Detection Theory (SDT) measures of performance (DePauli, 1967; Macmillan & Creelman, 2005). SDT is a theoretical framework which may be applied to both detection and recognition, and is particularly useful in situations of perceptual uncertainty (Lynn & Barrett, 2014). In a detection task, subjects decide how likely the presence of a signal is, relative to noise. In a recognition task (when a signal is known to be present), subjects must decide how likely the presence of signal A is, over signal B (Swets, 1966, outlined in (DePauli, 1967)). SDT is a useful tool for analysis as it allows subjects' responses to be broken down into two measures which depend on different internal processes, allowing a better understanding of the observed behaviour, particularly in situations with ambiguity: sensitivity to the signal and response bias. Sensitivity refers to the ability to discriminate between signal and noise or between two signal types. Response bias concerns the subject's tendency to respond with one option over the other. Assessing performance accuracy confounds sensitivity and bias measures (Lynn & Barrett, 2014). Subjects who perform very differently in terms of these indices could have similar response accuracies and thus, accuracy is not a good representation of behaviour (Lynn et al., 2014; Lynn & Barrett, 2014).

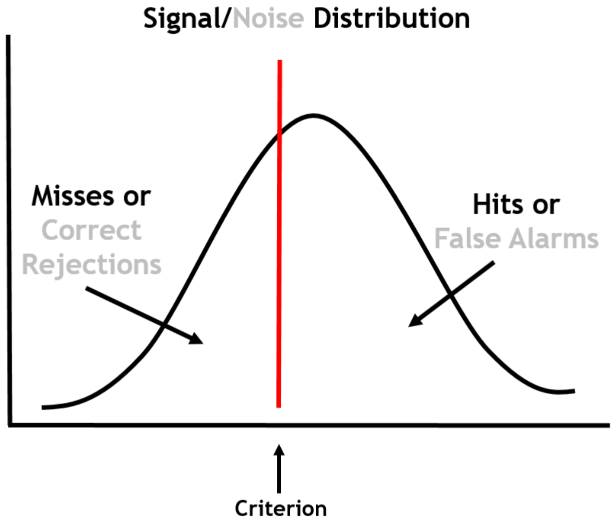


Figure 2.2 Distribution of signal and noise using SDT. The curve shows the distribution of signal or noise. If the curve denotes signal, then values above the criterion (*c*) result in hits and those below result in misses. If the curve represents noise, then responses above the criterion result in false alarms and those below lead to correct rejections. Criterion represents a subject's response criterion for responding in either direction.

To sum, we aim to look at how top-down and bottom-up signals interact to process degraded visual input. We will address this question using the following experiments which will utilise the partial occlusion paradigm outlined above:

- In the first and second experiments we will test whether consistent contextual information influences the detection and/or recognition of low contrast information. This will be explored by manipulating the consistency between the full-contrast surround and low-contrast target region and the visibility of the target regions. Participants will perform both a detection and recognition task to investigate whether the consistency between the high-contrast feedback signals and low-contrast feedforward signals influences these facets of perceptual processing.

- In the following experiment we will employ the same paradigm to test whether a consistency effect exists around the perceptual threshold. Here, the target region will be shown at contrast levels around the threshold for detection. The same tasks as the previous experiments will be employed.

- Next, we will explore whether a consistency effect extends beyond simple image features by testing whether congruency of the full-contrast surround can exert an effect even when simple image feature consistencies are removed.

- Finally, spatial frequency information will be limited within the stimuli to assess how limiting other image features can influence the effect of consistency in the recognition of degraded feedforward information.

Further details on the aim and methodology of each experiment can be found in section 2.3.6. Taken together, our results from this series of experiments should shed light on how the visual system uses top-down predictions to process upcoming sensory information. They will help allow us to determine how top-down predictions interact with the bottom-up signal and to what extent each of them contribute to the overall processing of the degraded information and how this manifests as a behavioural output.

2.3 Methods

2.3.1 Overall Rationale

The main aim of the project is to investigate how top-down and bottom-up information interact to influence the processing of dim feedforward input. In order to sufficiently address this, we look at the role of context and how the consistency between the contextual information (used to shape the top-down predictions) and the available feedforward information influences the processing of weak visual input. We also manipulate other aspects of the top-down signal as well as the bottom-up input and test subjects’ ability to both detect and interpret the degraded information. This is achieved through a series of psychophysical experiments. The general paradigm and procedure is outlined below, in addition to a summary of each experiment in turn.

2.3.2 Paradigm

All of the experiments followed the same general protocol but differed in aspects depending on the specific aim of each. Details are outlined for each experiment in turn in section 2.3.6. The basis of all the experiments, however, is an adaptation of a partial occlusion paradigm initially introduced by Smith & Muckli (Smith & Muckli, 2010).

The partial occlusion paradigm was developed to investigate the role of cortical feedback in early visual cortex. It works by occluding a portion of the stimulus image, eliminating feedforward input to the corresponding area of primary visual cortex. Subsequent activity in this corresponding cortical region provides evidence of feedback to primary visual cortex, in the absence of feedforward information. Here, we wanted to investigate whether these feedback signals could influence the processing of weak feedforward visual signals.

To achieve this, the occluded region in this series of experiments contained degraded (low contrast) feedforward information. We then manipulated the content of the feedback signals by altering the image in the un-occluded portion of the stimulus, to assess whether different contextual information influences the processing of the degraded information. The image in the un-occluded region (herein referred to as the 'surround' region), therefore either matched or did not match the low contrast 'occluded' information (herein referred to as the 'target' region). Figure 2.3 (a) depicts the location of the surround and target regions. To vary the degree of degraded information, we presented the target region at a variety of contrast levels. This enabled us to explore the moderating effect of context depending on the strength of the visual input.

2.3.3 Stimuli

Twenty-four scene images were selected from the Scene Classification Project Upright versus Inverted scene database (http://vision.stanford.edu/projects/sceneclassification/resources.html;Walther , Caddigan, Fei-Fei, & Beck, 2009). Split equally, the images depicted images from two basic-level categories (natural and manmade scenes) which in turn fall into six subordinate-level categories: beaches, forests, mountains, buildings, highways and industry scenes. We selected image exemplars from each category by applying the Spatial Envelope Model (Oliva & Torralba, 2001). The model ensures maximal dissimilarity between image features such as spatial frequency and orientation information while ensuring that this dissimilarity is equally represented within each category to eliminate the possibility that the variation could be biased in any categorical direction.

Once the images were selected, the stimuli were generated using MATLAB software version R2016a. Each stimulus image was comprised of two scene images: a full contrast surround image and a low-contrast target image. To create the images, the first image was selected from our twenty-four image stimulus database and grey-scaled. The second image was also greyscaled and the contrast reduced to one of the predetermined contrast levels. The two images were then merged so that the target image is shown in the bottom right corner and the surround image occupies the remaining three quadrants (as in Figure 2.3 (a)). This particular quadrant was chosen as the target region to replicate the stimulus layout of the stimulus used in previous partial occlusion paradigms within the lab (Muckli et al., 2015; Smith & Muckli, 2010). The stimulus image as a whole subtended 47.5° x 36.5° visual angle (VA), with the target region occupying 24.8° x 18.7° and the fixation cross the central 0.1° x 0.1°.

We manipulated the consistency between the surround and target regions to give rise to different task conditions, referred to as consistency level. These conditions are outlined in Figure 2.3(b). The Consistent Image condition contained a stimulus image in which the surround and target images both come from the same scene image. Thus, a Consistent Image stimulus looked like a complete scene, but with the bottom right quadrant being shown in low contrast. A Consistent Subordinate (level) stimulus had a surround and target image from within the same subordinate category but not from the same scene image (e.g., a beach and another beach image). A Consistent Basic (level) stimulus had a surround and target from within the same basic-level category but not the same subordinate-level category (e.g., a beach and a mountain; both are natural scenes but are not the same type of scene, nor the same image). An Inconsistent stimulus contained a surround and target from different basic-level categories (i.e., a natural surround and manmade target region or vice versa, for example, a beach and a building). Finally, a Target Only stimulus did not contain an image in the surround, only a low-contrast target image, eliminating contextual information in the surround. Collectively, these conditions were used across experiments, however the number and specific consistency conditions varied between experiments, details of which are outlined in section 2.3.6.

In addition to consistency between the surround and target being manipulated, we presented the target region at a range of low contrast levels, thus varying the degree to which the feedforward information was degraded. It is important to note however, that all the contrast levels are very low, meaning the target region was always substantially degraded compared to the full-contrast surround. Depending on which experiment, contrast levels ranged from either 0%-6.4% or 0.6%-1.8%. Figure 2.3 (c) gives example stimuli at each contrast level used.

2.3.4 Subjects & Procedure

We recruited subjects via an online subject pool through the University of Glasgow Institute of Neuroscience and Psychology. Recruitment stipulations were healthy (or corrected-to-healthy) vision and that subjects have not already taken part in a previous experiment from this project series, as each experiment was run consecutively. They attended one testing session which took place at the University's Institute of Neuroscience & Psychology. Before participation in the experiment, we verbally explained details of the task to each participant, and they were provided with an information sheet to read through before starting the experiment. We obtained informed consent prior to the experiment commencing in addition to demographic details for reporting. Subjects were paid GBP 6.00 per hour for their participation. Data for subjects were excluded on the basis of them not understanding task requirements or incompletion of the task.

During the testing session, subjects were seated 45cm from a desktop computer monitor (OptiPlex 9030 AIO with a 24" 1080p hd screen), replicating the distance within the 3T scanner in the Institute. Head position was supported with a chin rest. Screen resolution was set to 1024 x 768 pixels, replicating that of the in-scanner projector and screen brightness was kept at the default setting.

We ran each experiment using Presentation software (Neurobehavioral Systems, Inc., Berkeley, CA; www.neurobs.com). Although individual experiments varied in terms of stimulus type and number, each consisted of four experimental runs. Task instructions were displayed at the beginning of each run where subjects were required to press a key in order to continue, to allow enough time to read the instructions and ask any further questions. For each trial, the stimulus image was flashed at 4Hz (on/off 4x) for 1000ms, interleaved with a blank grey full-screen image containing the central fixation cross to maintain fixation. Flickering stimuli were chosen with the prospect of a future fMRI experiment. V1 is known to respond to stimulus flicker (Sekuler et al., 2002) and this rate is used is previous work within the lab (not yet published). Eye movements were not monitored within this series of experiments due to limitations with the availability of eye tracking equipment for psychophysical experiments, however the importance of eye tracking within visual perception studies (see (Saito & Sadoshima, 2016) for a review) is noted and is a measurement to consider for future experiments following the same principles.

After the stimulus presentation, the task question was presented for a maximum duration of 5000ms, but the next trial commenced as soon as the subject responded. Subjects indicated their response by pressing one of two response keys on the computer keyboard. The trial process is depicted in Figure 2.3 (d). Between runs, subjects were permitted take a break for as long as needed to maintain comfort throughout the experiment. Each run followed the same protocol and used the same stimulus set; however, stimulus order is pseudorandomised and differed between runs. Table 2.3 provides a summary of each Experiment and the aspects in which they differ.

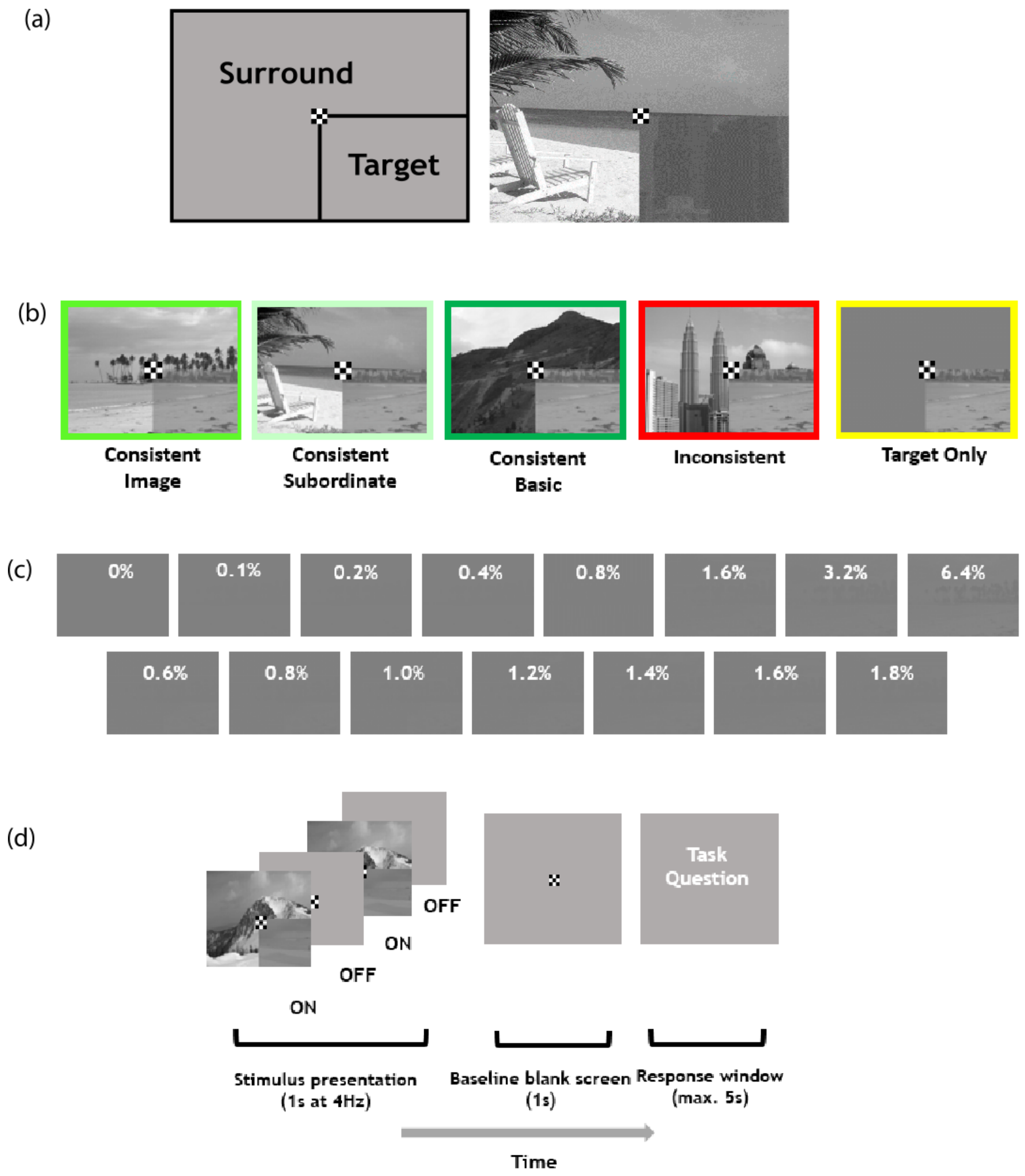


Figure 2.3 Psychophysical experiment details. (a) Depiction of the surround and target region locations (left) and example stimuli showing these two regions (right). Note that the contrast level of the target region on the right is higher than that used in the experiment, for demonstration purposes. (b) Stimulus conditions. Here, the target regions are shown at full contrast to illustrate the images used. Consistent image refers to a stimulus in which the surround and target regions contain the same image; Consistent Subordinate refers to a stimulus in which the surround and target regions contain an image from the same subordinate-level category but not the same image (e.g. a beach and another beach); Consistent Basic – the surround and target regions contain images from the same basic-level category (natural or manmade) but not the same subordinate-level category (e.g. beach and mountain); Inconsistent refers to a stimulus in which the surround and target region scenes differ at the basic level (e.g. a natural surround and manmade target region); Target Only – a stimulus in which there is no information in the surround. (c) Illustration of the contrast levels used in the experiments. Experiments 1 & 2 used contrast levels ranging from 0% to 6.4% (top row) whereas The Threshold Contrast Experiment onwards used contrast levels between 0.6% and 1.8% (bottom row). (d) Depiction of a trial, starting with flickering stimulus presentation and ending with the task question.

2.3.5 Analysis

Analysis for all experiments was be performed in both MATLAB (R2016a) and IBM SPSS Statistics (v27). We calculated participant detection rate and performance accuracy in the identification task on an individual basis and then averaged for a group-level analysis. We then tested the effects of consistency and contrast (and their interaction) on detection rate and identification accuracy using a repeated measures ANOVA for each experiment and conducted post-hoc tests using Bonferroni-correction. The dependent variable was performance accuracy (or detection frequency) and the independent variables were contrast (with 7 or 8 levels depending on which experiment) and consistency condition (ranging from 3-5 levels depending on experiment). We also analysed the data using a Signal Detection framework to give a measure of sensitivity and response bias with criterion and d' serving as the dependent variables and independent variables again being contrast level and consistency condition.

To analyse the data using Signal Detection Theory (SDT), we split the responses into hits, misses, false alarms and correct rejections (H, M, FA, CR) in two different ways. In the first instance, used the content of the target region along with the subject responses to determine H, M, FA and CR rates (Table 2.1). Secondly, we split the data according to the consistency between the surround and target regions, along with subject responses. Table 2.2 summarises how H, M, FA and CR rates were determined in this way.

We used the SDT analysis to give a measure of sensitivity (D prime (or d')) and criterion (response bias). Sensitivity refers to the ability for a subject to determine signal A over signal B and criterion provides a measure of how likely they are to respond a certain way. We considered these measures both overall and within each consistency condition.

SDT measures are calculated using a hit ratio and false alarm ratio. In the first analysis type (using only the content of the target region to determine H, M, FA and CR), hit rate was calculated as the number of hits/the number of trials in which the target contained a natural scene (or a beach, depending on which experiment), FA rate was calculated as the number of FA/the number of trials in which the target was manmade (or a mountain). In the second SDT analysis, where H, M, FA and CR was determined using the consistency between the surround and target regions, hit rate will be determined as the number of hits/the number of trials in which the surround and target were consistent. FA rate was calculated as the number of FA/the number of trials in which surround, and target were inconsistent.

<table>
<tr><th colspan="4">Experiments using natural and manmade stimuli</th></tr>
<tr><td></td><td colspan="3">Target Stimulus</td></tr>
<tr><td rowspan="3">Response</td><td></td><td>Natural</td><td>Manmade</td></tr>
<tr><td>Natural</td><td><i>Hit</i></td><td><i>False Alarm</i></td></tr>
<tr><td>Manmade</td><td><i>Miss</i></td><td><i>Correct Rejection</i></td></tr>
<tr><td colspan="4"></td></tr>
<tr><td colspan="4">Experiments using beach and mountain stimuli</td></tr>
<tr><td></td><td colspan="3">Target Stimulus</td></tr>
<tr><td rowspan="3">Response</td><td></td><td>Beach</td><td>Mountain</td></tr>
<tr><td>Beach</td><td><i>Hit</i></td><td><i>False Alarm</i></td></tr>
<tr><td>Mountain</td><td><i>Miss</i></td><td><i>Correct Rejection</i></td></tr>
</table>

Table 2.1 Calculation of H, M, FA and CR using the content of the target region. Here, these categories are determined using the content of the target region only and how this relates to the subject response. For example, a natural target stimulus and natural response results in a *hit* whereas a natural target stimulus and manmade response is categorised as a *miss.* Categorisation is shown for both stimulus types (natural/manmade (top) and beach/mountain (bottom)).

Experiments using natural and manmade stimuli			
Surround	**Target**	**Response**	**Category**
Natural	Natural	Natural	*Hit*
Manmade	Manmade	Manmade	
Natural	Natural	Manmade	*Miss*
Manmade	Manmade	Natural	
Natural	Manmade	Natural	*False Alarm*
Manmade	Natural	Manmade	
Natural	Manmade	Manmade	*Correct Rejection*
Manmade	Natural	Natural	
Experiments using beach and mountain stimuli			
Surround	**Target**	**Response**	**Category**
Beach	Beach	Beach	*Hit*
Mountain	Mountain	Mountain	
Beach	Beach	Mountain	*Miss*
Mountain	Mountain	Beach	
Beach	Mountain	Beach	*False Alarm*
Mountain	Beach	Mountain	
Beach	Mountain	Mountain	*Correct Rejection*
Mountain	Beach	Beach	

Table 2.2 Calculation of H, M, Fa and CR when the consistency between the surround and target regions is considered. Here, subject responses are categorised based on their ability to determine whether the surround and target region are consistent. Consistency between the surround and target regions either results in hits or misses, depending on the subject response, whereas inconsistency between the two regions results in either a false alarm or correct rejection.

2.3.6 Rationale for and details of each experiment

This project consists of a series of psychophysical experiments. Together they aim to aid our understanding of how top-down and bottom-up inputs interact to process degraded visual information. Separately, the experiments explore different aspects pertaining to this central question. The rationale for each experiment is outlined below, each aims to either fully or partly address a specific research question.

2.3.6.1 Does consistent context influence the detection and/or recognition of low-contrast information?

This section is comprised of results from two experiments (the Consistency Effect Experiment and Simplified Experiment). In the Consistency Effect Experiment (N=22, Mean Age=21.86, SD=2.27, 20 female), we aimed to test how different levels of consistency within the context influences the processing of degraded feedforward visual input.

The consistency between the context and low-contrast information was manipulated to several degrees, meaning the identity of the low-contrast scene could be determined by either simple image features such as line continuation or higher-level abstract visual features represented within the scene category. This experiment employed all of the twenty-four scene images from two basic-level categories: natural and manmade scenes. Each of these basic categories contained three subordinate-level categories, each with four image exemplars. The natural scene category contained (the subordinate categories depicting) images of beaches, mountains and forests and the manmade scenes were buildings, highways and industry. The stimulus images were comprised of two of these images combined depending on the consistency condition it belonged to. Experiment 1 consisted of five experimental conditions: Consistent Image (containing the same image in the surround and target regions), Consistent Subordinate (both regions contained images from the same subordinate category, e.g. beach and another beach), Consistent Basic (both regions contained images from the same basic level category but different subordinate categories, e.g. beach and mountain), Inconsistent (both regions were inconsistent at the basic-category level, e.g. beach and building) and Target Only (in which no information was depicted in the surround). Including such a range of consistency levels aimed to help distinguish which level of categorisation predictions operate at when processing degraded scene information; whether it can only exert an influence when low-level features are consistent or if more abstract consistencies are sufficient to drive an effect.

The target region in each of these consistency conditions was shown at 8 different contrast levels: 0% (no image shown in the target region), 0.1%, 0.2%, 0.4%, 0.8%, 1.6%, 3.2% and 6.4%. These contrast levels were chosen to encompass the perceptual threshold which is considered to be around 1% (Pelli & Bex, 2013). Per run, there were six repetitions of each condition at each contrast level, counterbalanced to ensure equal representation across subordinate categories.

The Consistency Effect Experiment followed a within-subjects design in which subjects took part in both a detection and recognition task in every trial. After each stimulus image, subjects were firstly presented with the recognition task where they were asked to report whether the target region contained a natural or manmade scene. This was then followed by a detection task in which we asked subjects whether they thought there was a scene present in the target region. This task order was deemed more intuitive than firstly asking whether or not a scene was present and then asking subjects to identify the scene, regardless of their response to the previous question. By employing a within-subjects design in this experiment, we were able to gauge both detection and recognition ability for the same stimuli, within the same subjects. We were also able to look separately at recognition-task performance in trials that were detected or undetected.

In the Simplified Experiment, we streamlined the design to allow for a more focused investigation and to test the effects of context on detection and recognition separately. In this experiment, the context and low-contrast portions of the visual scene were either consistent or inconsistent in the most explicit sense to test the rudimentary effect of contextual congruency on the processing of degraded information. Therefore, we used only two natural scene images (i.e., from the same basic-level category) to generate the stimuli: one of a beach and one of a mountain. This results in only three consistency conditions: Consistent (termed Consistent Image in the previous experiment but herein simply referred to as ‘Consistent’), Inconsistent and Target Only. The stimulus images were presented at the same eight contrast levels as Experiment 1, ranging from 0%-6.4%. Between-subject effects of detection (N=7, Mean Age=22.86, SD=3.58, all female) and recognition (N=7, Mean Age=28.43, SD=6.48, 3 female) were tested separately (contrasting the within-subjects design of the previous experiment) as we reasoned this would eliminate any task-related confusion which may arise from performing both tasks together (for example, being asked to identify a stimulus which has been explicitly confirmed as not been detected), thus minimising any potential confounding effects this may have. As there is no basic category-level distinction between the two stimulus images, the recognition task question in this experiment asked subjects if the target image contained a beach, with a yes/no response option.

2.3.6.2 Does a consistency effect persist around threshold contrast level?

The aim of the next experiment (the Threshold Contrast Experiment) was to focus on the perceptual threshold to see whether a consistency effect exists when feedforward information is degraded as such that it is only just perceptually visible. It followed the simplified design of two natural scene images (beach and mountain) and three consistency conditions (Consistent, Inconsistent and Target Only). Instead of contrast levels ranging from 0%-6.4% as in the first two experiments, the target region stimuli in The Threshold Contrast Experiment were shown at the following seven contrast levels: 0.6%, 0.8%, 1.0%, 1.2%, 1.4%, 1.6% and 1.8%. These levels were established a-priori based on the detection task results from the previous experiments and were identified as encapsulating the perceptual threshold across participants. Again, we tested both detection (N=10, Mean Age=24.20, SD=3.33, 9 female) and recognition (N=10, Mean Age=27.20, SD=7.55, 6 female) separately in a between-subjects design, using the same tasks as the previous experiment. If top-down predictions enhance or amplify consistent information, then we would expect a beneficiary effect even around threshold level.

2.3.6.3 Does a consistency effect extend beyond simple image features?

The next experiment termed Beyond Simple Image Features Experiment (N=28, Mean Age=22.14, SD=3.64, 20 female) aimed to look at whether a consistency effect can extend beyond the simple image feature of line continuation. In other words, do we see a facilitatory effect of consistency even when the context and target scenes are not part of the same image? If so, such an effect cannot be attributed to line continuation from the high contrast surround into the low contrast target region.

To test this, we used consistency conditions replicating those from the first (Consistency Effect) experiment whereby surround and target matched either on the Subordinate- or Basic- category level but did not include a Consistent Image condition. The contextual information was instead consistent in terms of higher-level (task relevant) information. Stimuli in the Inconsistent condition did not match at the basic category level and again, we also included a Target Only condition. Contrast levels ranged from 0.6%-1.8% (i.e., around perceptual threshold) and subjects were asked to indicate whether the target image was natural or manmade (in a recognition task). This experiment did not include a detection task. A consistency effect in this experiment would imply that predictable information influences the processing of degraded visual information beyond the effects of simple image features.

To further address this question, a separate experiment (N=29) was run via the online platform Pavlovia (www.pavlovia.org), with subjects recruited using Prolific (www.prolific.co). This experiment investigated whether a consistency effect extends beyond simple image features when the possibility of a response bias is controlled for a-priori, thus we term this experiment the Response Bias Experiment. It is important to note that I was not directly involved in the running of this particular experiment as it was undertaken as part of an MSc project. However, it adds value to the experimental series by addressing the key point of a potential response bias and therefore is included in this chapter with permission from those directly involved.

In the Response Bias Experiment, the opportunity for subjects to respond according only to the information in the context was eliminated, and instead they were provided with a task in which the content of both the context and the low-contrast scene information needed to be taken into account. This allowed us to determine whether our results so far reflect a perceptual process attributed to top-down influence (enhanced processing of consistent information) or a bias to overcome the task difficulty (a tendency to base responses on information that is more readily available).

This experiment shares the same underlying principle (of partial occlusion) as the previous experiments but does differ in a few critical aspects to allow response bias to be sufficiently addressed. Here, stimulus images are comprised of three rather than two sections. The top half of the image was shown at full contrast and was termed the context (akin to the surround in the previous experiments). The image in the context was either a beach, a mountain, a forest or left as a blank grey image. The bottom two quadrants contained two different scenes, each of which was shown at low contrast (6 levels: 0.3%, 0.6%, 1.2%, 2.4%, 4.8%, 9.6%). The images in these quadrants were always a beach and a mountain but the order of which was randomised, i.e., sometimes the beach was on the left and sometimes the right.

Whilst maintaining central fixation, subjects were asked to identify which image (left or right) depicts a specific type of scene (e.g., ‘which image was a beach?’) and indicate their result using the left and right arrow keys. Depending on the type of image outlined in the task (some subjects were asked to identify the beach scene and others the mountain scene, however, the requested scene was maintained for each subject throughout the whole experiment). The type of scene referred to in the task question is termed the target and the other scene the non-target. There were six stimulus images used (2 of each beaches, forests and mountains) and also six experimental conditions, which arise from the consistency between the context and target versus non-target regions. 1 – same target: context is consistent with the target image and is also the same image exemplar; 2 – different target: context is consistent with the target in terms of category but contains a different image exemplar; 3 – same non-target: context is the same image as the non-target region; 4 – different non-target: context is the same category as the non-target region but is a different image exemplar; 5 – inconsistent: context contains a forest and therefore does not match either the target or non-target region; 6 – no context: context region contains a blank grey image and therefore does not contain any contextual cues. Figure 2.4 summarises the experimental conditions used as well as the trial sequence presented to participants.

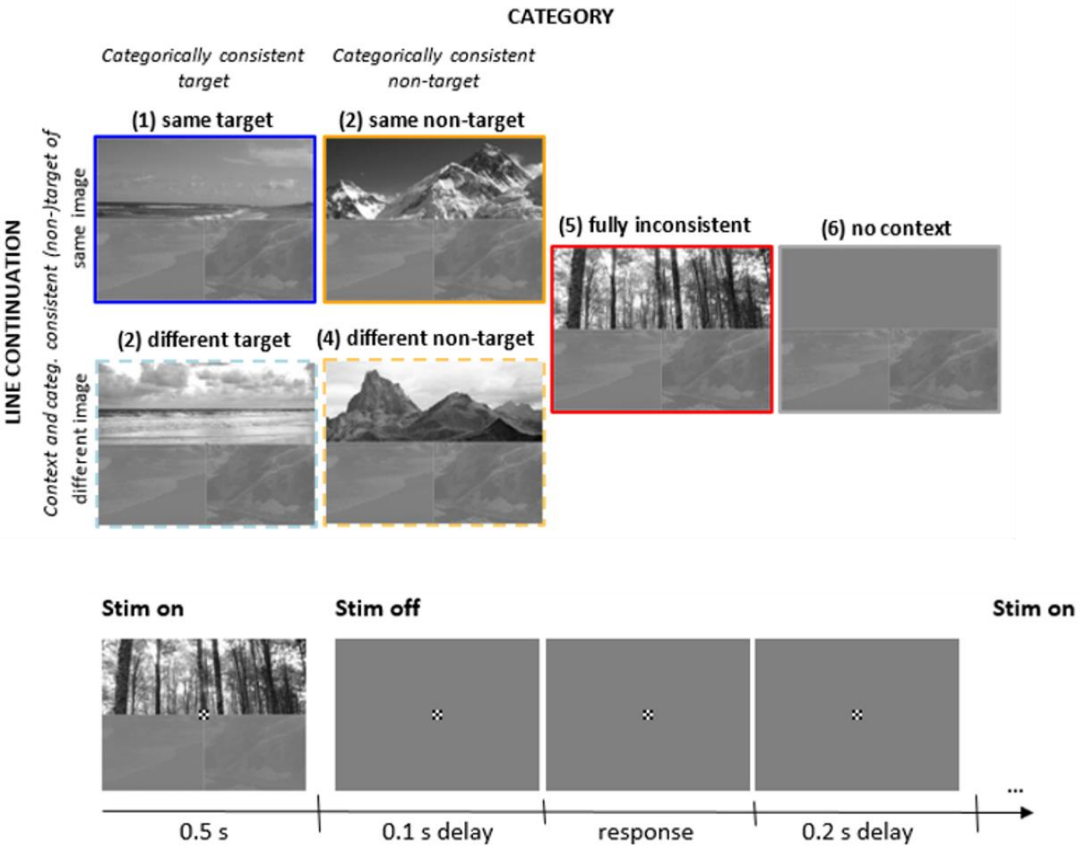


Figure 2.4 Response Experiment task details. The top portion of the figure outlines the experimental conditions if the task question asked subjects to indicate which region contained a *beach* (therefore making beach the target image). Same target – the context contains the same image as the target; Same non-target – the context contains the same image as the non-target region; Different target – the context contains an image from the same category as the target but uses a different image exemplar; Different non-target – the context contains an image from the same category as the non-target region but uses a different image exemplar; Fully inconsistent – the context contains a forest scene which does not match either the target or non-target regions; No context – the context contains a uniform grey image, providing no contextual information.

The bottom portion of the image depicts the trial sequence in which the stimulus is presented followed by a short delay before subjects are able to indicate their response via button key press.

Note that the stimulus images here are shown at higher contrast and the fixation cross is larger than in the experiment, for illustration purposes. Figures obtained with permission from the MSc student who ran the experiment.

Data were analysed using R (RStudio Team, 2019). A general linear mixed model (GLMM) was determined for each condition which fitted a logistic psychometric function to subjects’ accuracy (ability to determine the target stimulus correctly within each condition) as a function of contrast level. Performance was compared between conditions using a t-test at the 25%, 50% and 75% points on each curve. This allowed us to assess how well subjects could determine the low-contras scene information when the surrounding context differed in terms of its consistency.

2.3.6.4 What happens if the spatial frequencies are limited?

In the next experiment, we aimed to explore the influence of different aspects of the top-down and bottom-up signals in order to investigate the type of information which may drive a consistency effect. To do this, we manipulated the spatial frequency information in both the surround and target regions, altering their low-level properties. This experiment is therefore referred to as the Spatial Frequency Experiment.

We ran two separate experiments limiting the spatial frequencies in either the surround or target region (namely the Spatial Frequency Target Experiment and Spatial Frequency Surround Experiment, in turn). In the first (Spatial Frequency Target Experiment, (N=10, Mean Age=30.3, SD=10.10. 8 female)), we altered the spatial frequency information in the target region and not in the surround. The target region was therefore shown at low contrast and with limited spatial frequencies (low versus high), further degrading the target information. The spatial frequencies in the surround were maintained, thus the contextual information remained intact. We followed our standard paradigm of three main conditions (Consistent (same image), Inconsistent and Target Only) and seven contrast levels (0.6%-1.8%), however, we introduced the additional variable of spatial frequency so that each consistency condition was shown at both a low (only, LSF) and high (only, HSF) spatial frequency (with an equal number of trials in each condition). Spatial frequency of the stimulus images was altered using the MATLAB function Butterfilter which applies a low- or high-pass filter to generate low and high spatial frequency stimulus images. Figure 2.5 gives example low and high spatial frequency stimulus images. Subjects performed a recognition task only, in which they were asked whether the target region contained a beach.

In the second experiment (Spatial Frequency Surround Experiment, (N=10, Mean Age=25.3, SD=5.48, 5 female)) we manipulated the spatial frequency information in the surround only. Thus, testing how limited spatial frequencies available in the context influences the processing of degraded feedforward information. In a sense, this version of the experiment tests how degraded contextual information influences the processing of degraded feedforward information. The target region was again shown at low contrast, but its spatial frequency was not altered.

We expect to find differences in recognition abilities when contextual (surround) information is limited to either high or low frequency domains. High spatial frequency context will provide information about edges and will therefore still enable processes exploiting these image features (such as line continuation) to be used. Low spatial frequency information on the other hand provides a more coherent representation of the scene as many visual features can still be distinguished when high spatial frequency information is removed but may prompt a more abstract utilisation of the top-down predictions. This may or may not differentially influence subjects' ability to determine the low-contrast target. Conversely, limiting spatial frequencies in the target region further degrades the available signal, therefore we may not find any substantial differences between spatial frequency types in the target region if the task becomes too difficult as a result. If the task is still manageable with this limitation imposed, then we may find that subjects are better at identifying the low-contrast scene when it is shown at low (versus high) spatial frequencies due to previous findings within the literature showing that the gist of a scene can be gleaned from low spatial frequency information.

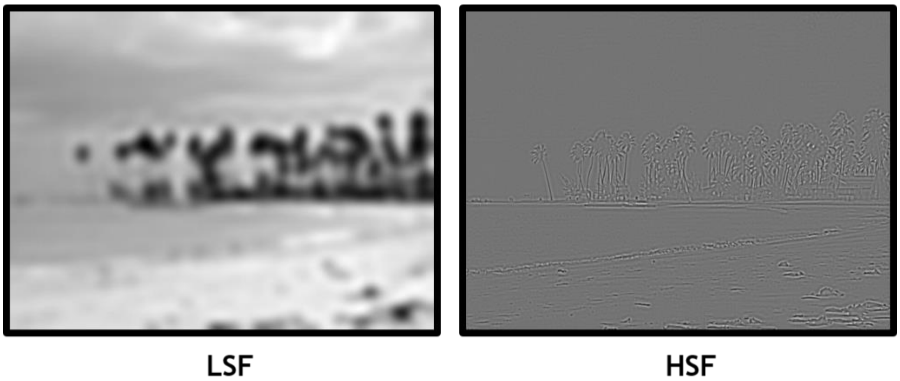


Figure 2.5 Illustration of low and high spatial frequency information. Stimulus image depicting the same beach scene, with the spatial frequency information limited to low spatial frequencies on the left and high spatial frequencies on the right. Low spatial frequencies capture the gist of the scene whereas high spatial frequencies convey fine-scale information about boundaries and edges.

2.3.7 Summary

Through this series of experiments, we explored different characteristics and aspects of predictive processing within the visual system to investigate how this processing system operates within the realm of vision. More specifically, we addressed our fundamental experimental aim of investigating how top-down and bottom-up inputs interact to influence the processing of degraded visual information.

Experiment Name	N	Task	Rationale
Consistency Effect Experiment	22 (within)	Detection Recognition	To test for a consistency effect across 5 levels of categorical consistency, using 8 contrast levels ranging from 0%-6.4%.
Simplified Experiment	7 7	Detection Recognition	Simplifying the design to include only 3 consistency levels and the same 8 contrast levels. Detection and recognition tested separately.
Threshold Contrast Experiment	10 10	Detection Recognition	Testing for a consistency effect around the threshold (0.6%-1.8% contrast). 3 levels of consistency.
Beyond Simple Image Features Experiment	28	Recognition	Testing whether an effect extends beyond the simple image feature of line continuation. 4 levels of consistency (excluding consistent image). Contrast 0.6%-1.8%.
Response Bias Experiment	29	Recognition	Testing whether a consistency effect exists beyond line continuation when a response bias is accounted for. Used 3 images to constitute the stimulus image. Contrast levels 0.3%-9.6%.
Spatial Frequency Target Experiment	10	Recognition	Testing whether limiting spatial frequencies in the target region influences recognition. Contrast levels 0.6%-1.8%.
Spatial Frequency Surround Experiment	10	Recognition	Testing whether limiting spatial frequencies in the surround influences recognition. Contrast levels 0.6%-1.8%.

Table 2.3 Breakdown of individual experiments. Outline of key experiment details; number of subjects, type of task involved and a summary of the main rationale of each.

2.4 Results

2.4.1 Consistent context enhances recognition (but not detection) of low-contrast information

We aimed to test whether consistent contextual information influences the perception of low-contrast feedforward input. We found that reliable information in the surround enhances the recognition of low contrast information but does not improve the ability to detect it. We ran a repeated measures ANOVA with detection frequency being the dependent variable and the independent variables being contrast level (7 or 8 levels, depending on the experiment) and consistency condition (5 or 3 levels depending on the experiment). Details of our findings from two experiments demonstrating this are outlined below.

2.4.1.1 Detection Frequency

We ran a repeated measures ANOVA to examine the effects of consistency (between the surround and target regions) and contrast level on subjects' ability to detect low-contrast information in the target region. Results reported using Greenhouse-Geisser correction, showed no significant differences in subjects' detection frequency between conditions in either the Consistency Effect Experiment (F(1.414, 29.686)=1.553, p=.228) or The Simplified Experiment (F(1.120, 6.720)=.684, p=.453). Detection rates are similar across all conditions in each experiment, as shown in Figure 2.6 (top row), suggesting that consistency of contextual information does not influence the ability to detect low-contrast information.

We find a highly significant effect of contrast in both the Consistency Effect Experiment (F(1.600, 33.610)=217.422, p<.001), with a large effect size of η_p^2=.912 and The Simplified Experiment (F(1.533, 9.199)=154.301, p<.001), also with a large effect size of η_p^2=.963. As expected, subjects' ability to detect low-contrast information in the target region increases with contrast level. Post-hoc tests using Bonferroni correction confirm this effect (ps<.005) and reveal significant differences between lower and higher contrast levels. Visual inspection of these plots confirms the findings that detection improves with increased contrast level but is not influenced by the consistency between the high-contrast surround and low-contrast target information. We find a significant interaction for consistency x contrast in the Consistency Effect Experiment (F(8.821, 185.247)=2.795, p=.005), in which differences in detection rates between consistency conditions decrease as contrast levels increase. However, these differences at the lower contrast levels are non-significant and this significant interaction does not persist in The Simplified Experiment.

2.4.1.2 Recognition Accuracy

Subjects also completed a recognition task in which they were asked to identify the type of scene depicted in the low-contrast target region. The same subjects completed both the detection and recognition tasks in the Consistency Effect Experiment in a within-subjects design, whereas The Simplified Experiment followed a between-subjects design with different subjects taking part in each task. Using recognition accuracy as the dependent variable in the repeated-measures ANOVA, in the Consistency Effect Experiment we find a significant main effect of consistency (F(1.285, 26.988)=106.582, p<.001), reported using Greenhouse-Geisser corrections with a large effect size of η_p^2=.835. Post-hoc tests using Bonferroni correction reveal significant differences between all consistency conditions (ps<.005) except the Consistent Subordinate (e.g., beach and another beach) and Consistent Basic (consistent within the basic-level category of either natural or manmade, e.g., beach and mountain) conditions (p>.05), who have Mean accuracies of 85.933% (SE=1.936%) and 86.024% (SE=1.760%), respectively. The Consistent Image condition (in which the surround and target images are from the same image exemplar) provides the highest recognition accuracy (Mean=89.561%, SE=1.694%) and the Inconsistent condition results in the lowest performance accuracy (Mean=41.496%, SE=2.812%). When there is no information in the surround region, subjects perform with a Mean accuracy of 67.931% (SE=1.169%). Thus, subjects' ability to identify low-contrast information appears to be enhanced when high-contrast information in the surround is consistent. Performance is significantly lower when there is no information provided in the surround (Target Only condition) and is further decreased when information in the surround is inconsistent. This suggests the ability to identify low-contrast information is somewhat hindered by contextual information which is not congruent.

This consistency-effect, however, is not replicated in The Simplified Experiment. In this experiment, we find no significant difference between consistency conditions on the ability to recognise the scene in the low-contrast target region. The same pattern of results as the Consistency Effect Experiment are observed, with highest performance accuracy being found in the Consistent condition (Mean=67.694%, SE=4.540%), followed by the Target Only condition (Mean=63.974%, SE=4.153%), and lowest performance in the Inconsistent condition (Mean=62.619%, SE=2.418%). However, these differences are slight compared to those found in the Consistency Effect Experiment and fail to reach significance (F(2,12)=1.189, p=.338).

Again, we find a significant effect of contrast in both the Consistency Effect Experiment (F(3.735, 78.435)=210.579, p<.001)) and the Simplified Experiment F(1.436, 8.617)=13.914, p=.003), reported using Greenhouse-Geisser corrections and each with a large effect size (η_p^2 =.909 and η_p^2 =.699, respectively). These reflect differences between lower and higher contrast levels, with higher performance accuracy when contrast levels are higher (post-hoc pairwise comparisons reveal the distinction between lower and higher contrast levels to be highly significant in The Consistency Effect Experiment (ps<.001) and slightly less pronounced in The Simplified Experiment (ps<.05), implying that subjects found the task easier when the stimulus was more easily perceived.

Figure 2.6 (bottom row) reveals performance accuracy in both experiments for each consistency condition, at every contrast level. We see the clear consistency effect in the Consistency Effect Experiment, in which consistency between the surround and target regions improves performance compared to having no information in the surround. This consistency-effect appears to be bi-directional as inconsistency between the surround and target appears to hinder performance. This pattern can also be observed in The Simplified Experiment but does not reach significance. As can be seen by visual inspection of The Simplified Experiment, performance increases with contrast level until the highest contrast level (6.4%) whereby performance decreases slightly in all conditions. However, this is non-significant, (performance between contrast levels 3.2% and 6.4% do not differ significantly (p>.05)).

We also observe a significant interaction between consistency and contrast in the Consistency Effect Experiment (F(6.902, 144.946)=32.540, p<.001), which reflects a decrease in the consistency effect as contrast levels increase. That is, as the low-contrast information becomes more visible, subjects appear to be less influenced by the consistency of the high-contrast context. We do not see a significant interaction in The Simplified Experiment.

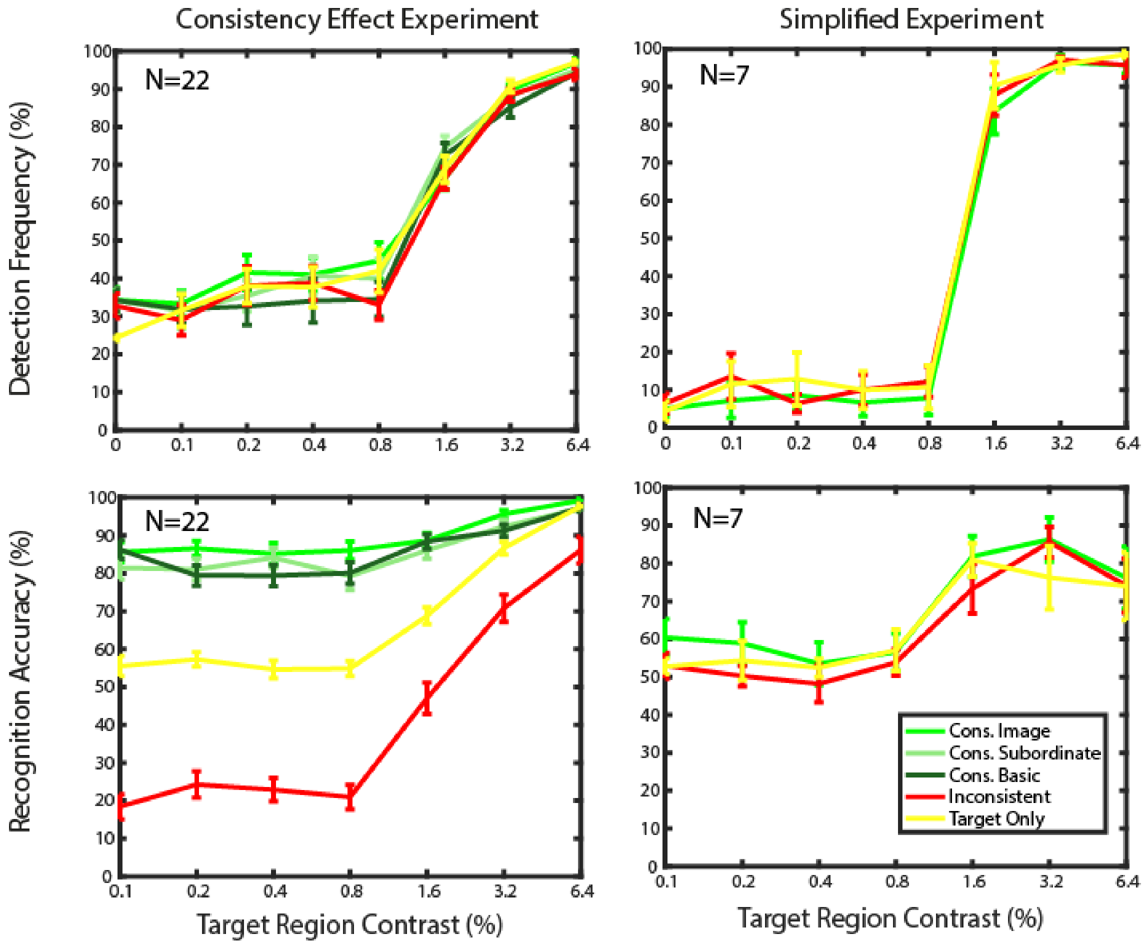


Figure 2.6 Detection frequency and recognition accuracy for the Consistency Effect Experiment (N=22) and the Simplified version (N=14). Results of The Consistency Effect Experiment (left) and the Simplified Experiment (right) for both the detection task (top row) and recognition task (bottom row). The Detection task results depict the frequency at which subjects detected a scene in the target region at each contrast level and in each consistency condition. Results of the recognition task depict subject accuracy in being able to determine the content of the target region, at each contrast level and within each consistency condition. Error bars represent Standard Error.

The significant consistency effect we observe in The Consistency Effect Experiment suggests that when top-down predictions (formed on the basis of the contextual information in the surround) are informative, the content of the low-contrast target region is more easily recognised. In contrast, when these predictions are uninformative, in the case where the surround and target region are inconsistent, subjects’ ability to recognise the low-contrast scene is reduced. Taken together, these results imply that contextual information strongly influences the behavioural response to degraded feedforward information. However, by simply looking at accuracy as an indication of performance we cannot determine whether there is a response bias.

2.4.1.3 Response Bias and Sensitivity

The aim of the recognition task is for subjects to determine the type of scene (natural/manmade; beach/mountain) in the target region. If contextual information in the surround influenced the perception of the low-contrast information in the target region then we would expect to see improved recognition performance when the context is consistent relative to inconsistent, which is indeed what we find. However, we could also find this pattern of results if subjects were responding only according to the information in the surround. That is, we would expect a similar pattern of performance accuracy if participants were not actually better at identifying the content of the target region but were instead biased towards responding that the information in the target area is consistent with information in the surround. A response bias in this case would also result in increased accuracy when the surround and target are consistent and decreased accuracy when they are Inconsistent, as well as chance-level accuracy when there is no information in the surround.

To account for response bias, we will now consider performance in the recognition task in terms of Signal Detection Theory. Signal Detection Theory provides a measure of response bias (known as criterion) indicating how inclined subjects are to respond in a certain manner, as well as a measure of sensitivity (akin to performance accuracy but not influenced by response bias) in the form of d prime (denoted as d’). We ran two types of signal detection analysis, resulting from splitting the data in two different ways. Further details of Signal Detection Theory and its application within this experiment series can be found in section 2.3.5.

When taking only the content of the target region into account, a negative response bias means that subjects are more likely to respond that the target is natural (in The Consistency Effect Experiment) or a beach (in the Simplified Experiment). Using a repeated-measures ANOVA with criterion and sensitivity as the dependent variables in turn here, we find a significant main effect of consistency on bias in the Consistency Effect Experiment ($F(4,84)=4.160$, $p<.005$), with a large effect size of $\eta_p^2=.165$, in which subjects are more likely to respond with ‘natural’ when the surround and target are consistent (post-hoc pairwise comparisons reveal significant differences between the Consistent Subordinate and Inconsistent ($p<.005$) and also Consistent Subordinate and Inconsistent ($p<.05$). Importantly, this does not mean that subjects are more likely to respond that the target is natural if the surround is natural as the consistent condition includes trials in which both the surround and target are manmade. We do not find this significant effect of surround type in the Simplified Experiment ($F(1.151, 6.908)=.270$, $p=.652$), reported here using Greenhouse-Geisser corrections. We find a significant effect of contrast in the Simplified Experiment ($F(6,36)=63.845$, $p<.001$) with a large effect size of $\eta_p^2=.914$, with subjects being more likely to respond with beach as the contrast increases (post-hoc pairwise comparisons again reveal significant differences between the lower and higher contrast levels ($ps<.05$). This contrast effect is not observed in the Consistency Effect Experiment ($F(3.135, 65.843)=.875$, $p=.515$). We find a significant interaction between consistency x contrast in both the Consistency Effect Experiment ($F(24,504)=1.894$, $p=.007$), (effect size $\eta_p^2=.083$), and the Simplified Experiment ($F(12,72)=2.759$, $p=.004$) with a large effect size $\eta_p^2=.315$). It is perhaps worth mentioning that while the span of bias is quite small in the Consistency Effect Experiment, in the Simplified Experiment, subjects are biased towards saying that the target is a mountain when the contrast is low and a beach when the contrast increases. This could suggest a qualitative difference in the categories or stimulus images used within The Simplified Experiment and between both experiments.

Despite the potential for bias, we find a significant main effect of consistency in the Consistency Effect Experiment using the non-biased sensitivity (d’) as a measure of performance ($F(1.1477, 31.022)=55.374$, $p<.001$, reported using Greenhouse-Geisser corrections), (effect size $\eta_p^2=.725$), but not in the Simplified Experiment ($F(2, 12)=1.678$, $p=.228$). In The Consistency Effect Experiment we find an increased sensitivity when the surround and target region are consistent, relative to both no or inconsistent information in the surround (post-hoc pairwise comparisons reveal these effects to be highly significant (ps<.001). There is a significant effect of contrast in both experiments, the Consistency Effect Experiment: F(6,126)=151.717, p<.001; the Simplified Experiment: F(6,36)=12.507, p<.001, each with large effect sizes (η_p^2 =.878 and η_p^2 =.676, respectively). Post-hoc pairwise comparisons using Bonferroni corrections again reveal mainly significant differences between the low and high contrast levels in the Consistency Effect Experiment (p<.001) but the pairwise comparisons in the Simplified Experiment do not reach significance despite an overall main effect of contrast. We also find a significant interaction between consistency and contrast in the Consistency Effect Experiment only (F(9.865, 207.155)=6.212, p<.001), (effect size η_p^2 =.228), whereby the overall effect of a consistent surround decreases as contrast increases. Figure 2.7 summarises the SDT measures for the two experiments.

When applying Signal Detection Theory using consistency between the target and surround to calculate bias and sensitivity (Figure 2.8), a negative response bias indicates that subjects are more likely to say that the target region is consistent with the surround. Here, we consider all trials grouped together, rather than between consistency conditions, in order to investigate these measures overall. We find a significant effect of contrast on response bias in the Consistency Effect Experiment (F(2.607, 54.757)=8.378, p<.001), effect size η_p^2 =.285, with subjects less likely to respond that the surround and target are consistent as contrast levels increase. At the lower contrast levels, subjects are more likely to respond that the surround and target region are consistent, a tendency which reduces as contrast increases (however these differences are only reflected in a few of the Bonferroni-corrected post-hoc comparisons, ps<.05). This suggests that when the target region is the most difficult to see, subjects may be more reliant upon the information in the surround. This finding is not replicated in the Simplified Experiment, (F(2.207, 13.243)=.629, p=.563), meaning criterion values remained stable across contrast levels. Subjects do not exhibit much of a response bias at any contrast level.

Looking at sensitivity, we find a significant effect of contrast in the Consistency Effect Experiment (F(2.758, 57.923)=117.399, p<.001), with a large effect size

η_p^2 =.848. Post-hoc pairwise comparisons reveal highly significant differences between the highest three contrast levels and all other contrasts (ps<.001). When contrast levels are low, sensitivity is around chance level, however, this increases with increased contrast. As sensitivity is a measure that is independent of bias, we can therefore assume that task performance does improve with increased contrast, and that this improvement is not due to a response bias towards consistency between the surround and target regions. We also find a significant main effect of contrast on sensitivity in the Simplified Experiment (F(1.333, 8.000)=6.887, p=.025) (reported here with Greenhouse-Geisser corrections), effect size is large η_p^2 =.534. Again, sensitivity increases with contrast, however post-hoc comparisons using Bonferroni corrections do not reflect this significant main effect (ps>.05).

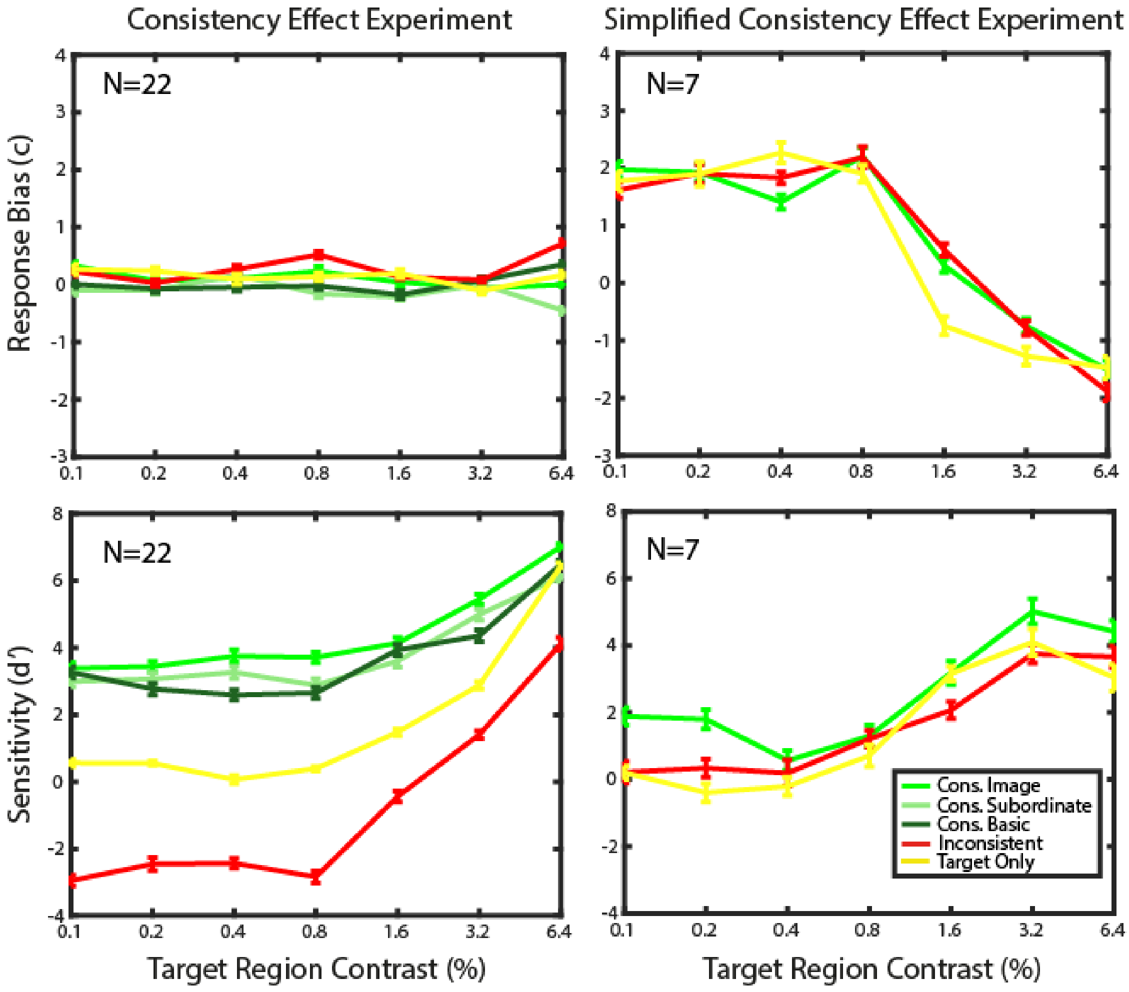


Figure 2.7 SDT results for Experiments 1 and 2. Response bias (top) and sensitivity (bottom) for each consistency condition, at each contrast level in the Consistency Effect Experiment (left) and the Simplified Experiment (right). Hits, misses, false alarms and correct rejections were determined using the content of the target region and subject response only. Error bars represent Standard Error.

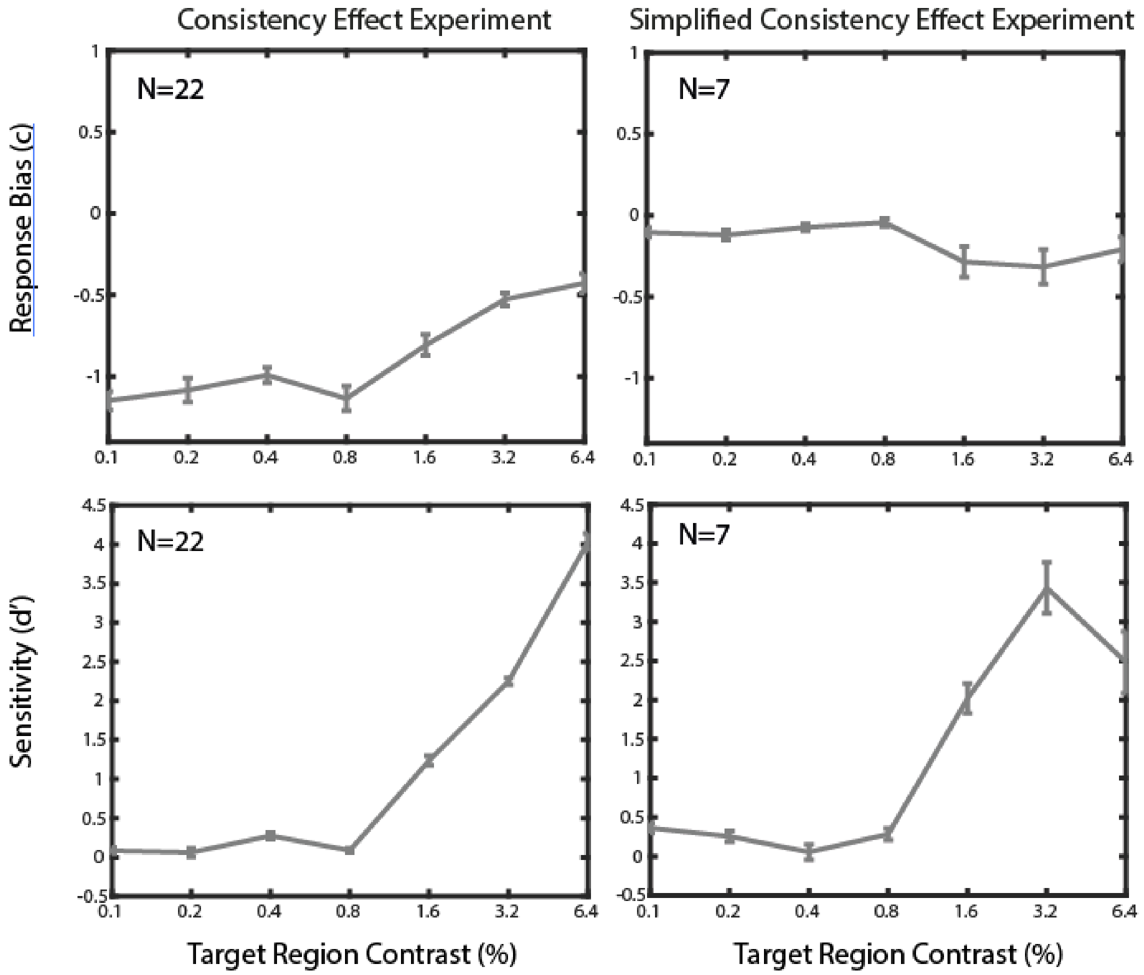


Figure 2.8 SDT results for Experiments 1 and 2 when consistency between surround and target is considered. Response bias (top) and sensitivity (bottom) results for Experiments 1 (left) and 2 (right) across all trials, for each contrast level. Hits, misses, false alarms and correct rejections here are determined using the consistency between the surround and target region in addition to subject response. Error bars represent Standard Error.

In sum, we do not find evidence to suggest that the consistency between the surround and target regions influences the ability to detect a low contrast scene. We do, however, find that different information in the surround influences the ability to identify the low-contrast scene. Performance accuracy improves with consistent information presented in the surround and decreases when inconsistent information is presented. This effect also manifests in sensitivity measures which are independent of response bias. Performance generally increases with increased contrast. It appears that when the contrast levels are low, subjects are more inclined to presume that the surround and target regions are consistent, but that their tendency to do so decreases with increased contrast, and therefore results in increased sensitivity when the target region is more visible.

2.4.1.4 Detected versus Undetected Trials

As the Consistency Effect Experiment was a within-subjects design, in which the same subjects were asked to perform both a detection and recognition task consecutively on each trial, we are able to investigate the effects of consistency and contrast on performance in the recognition task separately for detected and undetected trials. Detected trials refer to those in which the subject reported perceiving information in the low-contrast target region and undetected trials are those in which they reported not perceiving the low-contrast scene.

Although it seems counter-intuitive for subjects to perform a recognition task on trials which they report not being able to see a scene present, subjects were asked to give their best guess in a two-alternative recognition task, even if they felt like there was no stimulus present. Subjects were also always presented with the recognition task first on each trial, followed by the detection task, in an attempt to minimise the counter-intuitive nature of trying to identify a stimulus that one cannot perceive. By performing the recognition task first, the detection task in a sense provides a confidence rating of their recognition task response.

When focusing only on Detected Trials (56.66% of all trials), we find a significant effect of both consistency ($F(2.392, 45.441)=66.111$, $p<.001$) and contrast ($F(3.711, 70.509)=104.091$, $p<.001$) on performance accuracy, as well as a significant interaction between consistency and contrast ($F(7.326, 139.191)=9.188$, $p<.001$), each with large effect sizes ($\eta_p^2 =.777$, $\eta_p^2 =.846$, $\eta_p^2 =.326$, respectively). In terms of the effect of consistency, there are significant differences between all conditions (ps<.05 using Bonferroni corrections) except Consistent Basic category and the other two Consistent conditions (the difference between Consistent Image and Consistent Subordinate is significant ($p=.007$)). We see the same response pattern as when all trials are included, in that consistent conditions improve performance, relative to having no information in the surround, and performance decreases in the inconsistent condition.

Performance improves as contrast increases. Post-hoc pairwise comparisons using Bonferroni Correction reveal significant differences between the lower and higher contrast levels (ps<.001). Figure 2.9 shows performance accuracy for each consistency condition, at each contrast. As contrast increases, the differences in performance accuracy observed between consistencies decreases, indicating that the more easily a stimulus is perceived, the less influence context consistency has on performance.

Turning attention to the Undetected Trials only (i.e., trials in which subjects responded that they could not perceive a stimulus in the low-contrast target region), (41.71% of all trials) we again see the same pattern of results whereby there is a bi-directional consistency effect on performance accuracy. The consistency between the surround and target regions significantly influences performance on the recognition task (F(1.663)=105.650, p<.001) with a large effect size of $\eta_p{}^2$ =.834, with consistent contextual information resulting in the highest performance accuracy and inconsistent context resulting in the lowest recognition accuracy. Bonferroni-corrected post-hoc pairwise comparisons reveal a significant difference between the Inconsistent and Target Only conditions and all other conditions (ps<.001). While performance does increase with contrast level, this effect does not reach significance (F(2.016, 38.313)=2.865, p=.069), revealing less of an effect of contrast when the target is not detected.

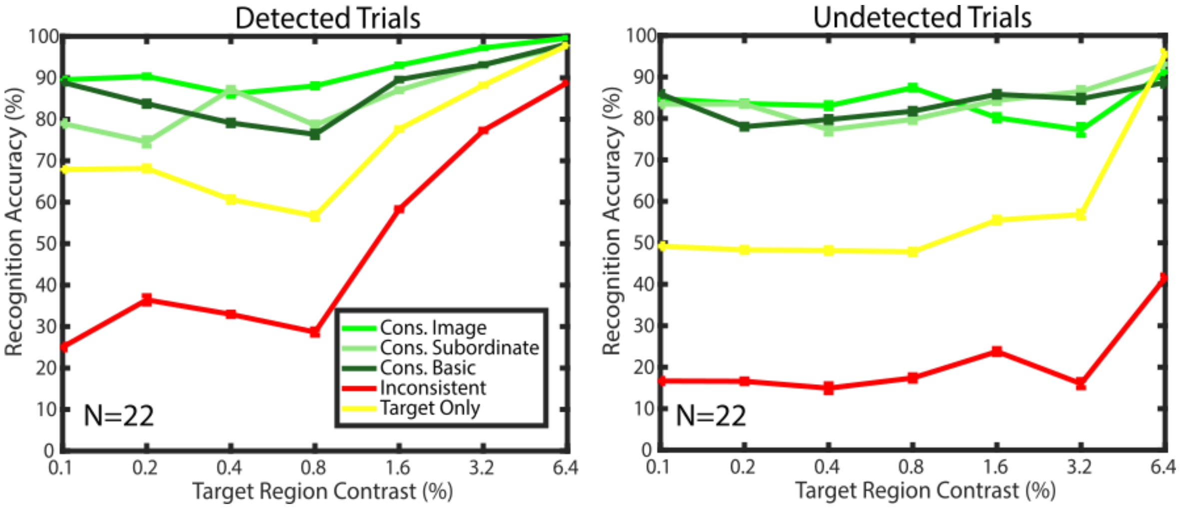


Figure 2.9 Recognition accuracy for detected and undetected trials. Results of the recognition task for The Consistency Effect Experiment for trials which subjects detected that there was a scene presented in the target region (left) and those in which they did not detect a scene (right). Recognition accuracy is plotted for each consistency condition, at each contrast level. Error bars represent Standard Error.

Applying the principles of SDT to our analysis firstly on detected trials only reveals no significant evidence of a response bias when a general bias towards responding in either direction (natural or manmade) is considered. Biases between conditions are all clustered close to zero and there are no significant differences between any of the conditions (ps>.05). We also find no significant effect of contrast, nor a significant interaction between condition x contrast.

In terms of sensitivity, we do see significant differences between conditions (F(1.386, 5.546)=10.559, p=.016), reported using Greenhouse-Geisser corrections, with a large effect size of η_p^2 =.725. The three consistent conditions (Consistent Image: Mean=5.082, SE=.534; Consistent Subordinate: Mean=3.972, SE=.477; Consistent Basic: Mean=3.685, SE=.548) resulting in the highest sensitivity rates. Sensitivity is higher than chance in the Target Only condition (Mean=3.173, SE=.195), implying subjects are still able to determine the content of the target region when no contextual information is present but not as effectively as when consistent contextual information is shown in the surround. Sensitivity is weakest in the Inconsistent condition (Mean=.464, SE=.739) with sensitivity being lower than chance at the lowest contrast levels. Bonferroni-corrected post-hoc comparisons reveal these differences to be non-significant. We find a significant main effect of contrast (F(2.592, 10.369)=19.853, p<.001) (effect size η_p^2 =.832), with sensitivity increasing with contrast level (post-hoc pairwise comparisons reveal significant effects between the highest contrast level and the lowest three contrast levels (ps<.05)). Although we find that sensitivity differences between the consistency conditions decrease with contrast, this interaction is insignificant (p>.05).

When testing for a response bias toward being inclined to respond that the surround and target are consistent, we do find evidence to suggest that subjects are responding in this way even in the detected trials. Criterion values are below zero for all contrast levels (although we do find a significant effect of contrast, with bias decreasing with increasing contrast (F(3.555, 74.651)=7.070, p<.001)), η_p^2 =.252, indicating a tendency to report that the surround and target are consistent, which differs significantly between the lowest and highest contrast levels according to post-hoc comparisons (ps<.05).

When this bias is controlled for however, we find that sensitivity levels are around chance at the lower contrast levels, implying that here subjects may be responding according to the surround. However, when the contrast levels increase, we find significantly improved sensitivity (F(4.023, 84.482)=52.544, p<.001), η_p^2 =.714, suggesting subjects are able to accurately determine the content of the target region at the higher contrast levels, independent of a response bias towards the surround. Post-hoc comparisons revel significant differences between the highest two contrast levels and all other levels (ps<.001). Figure 2.10 summarises the Signal Detection results for detected trials.

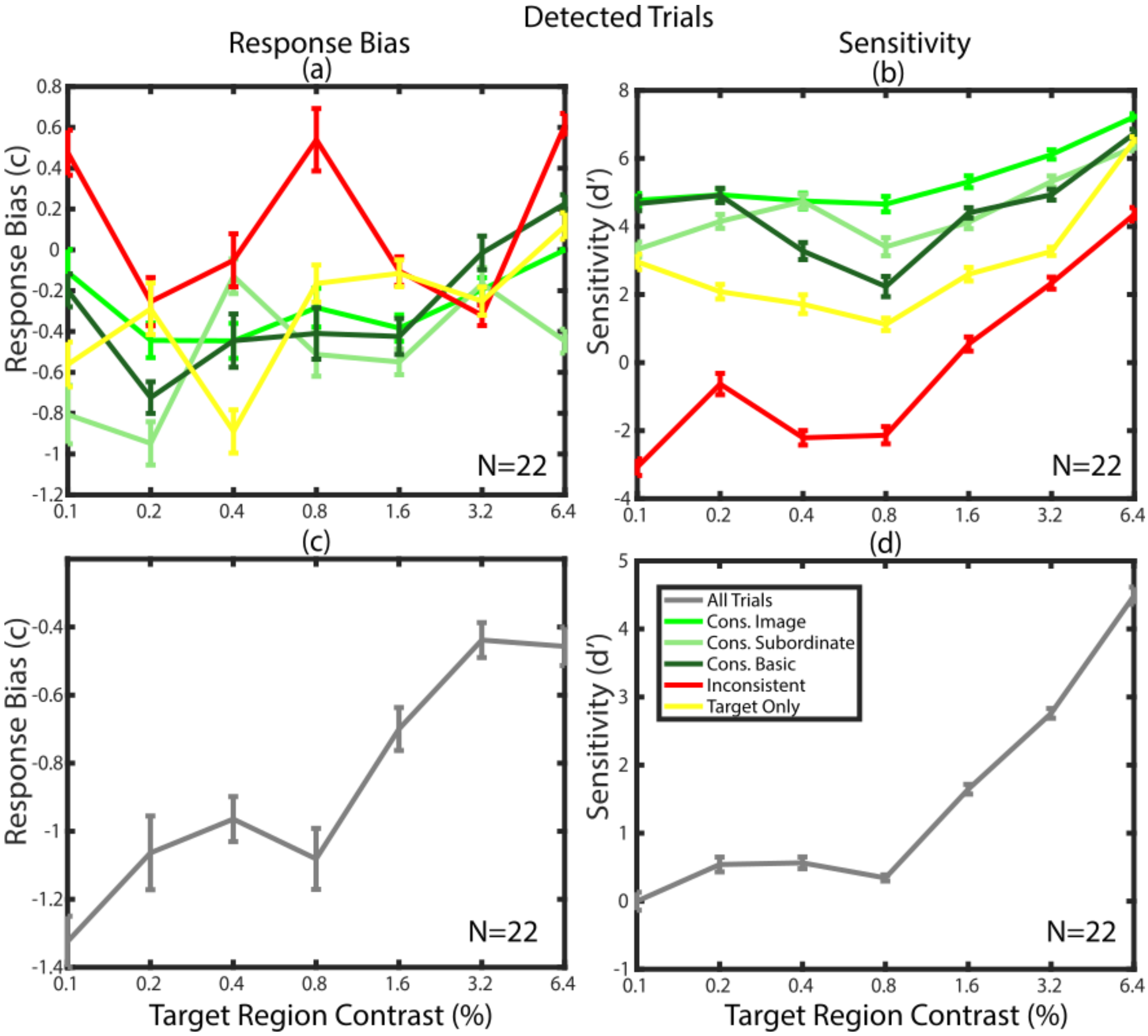


Figure 2.10 SDT results for detected trials only. Response bias and sensitivity measures for The Consistency Effect Experiment in detected trials only. The top row shows response bias (c) and sensitivity (d') in each consistency condition, at each contrast level. Here, hits, misses, false alarms and correct rejections are calculated using the content of the target region and the subject response. The bottom row shows overall response bias (left) and sensitivity (right) when hits, misses, false alarms and correct rejections are determined using the consistency between the surround and target regions. Error bars represent Standard Error.

When focusing only on Undetected Trials for our SDT-based analysis (Figure 2.11), we are not able to include the last contrast level in the analysis as there is an insufficient number of undetected trials at the highest contrast level. Therefore, analysis is performed on the first 6 contrast levels. In terms of a general response bias, we find no significant differences between conditions and biases are close to chance-level. We then observe the same pattern of sensitivity results, with increased sensitivity in the consistent conditions and decreased sensitivity in the Inconsistent condition, however, this difference does not reach significance. In fact, we find no significant main effect of either condition or contrast, nor a significant interaction between these factors for either response bias or sensitivity. This could reflect a lack of power due to low trial numbers available.

When considering a response bias toward the surround, the results are more telling. Again, we excluded the highest contrast level here due to lack of undetected data. We observe a negative response bias across all contrast levels, meaning subjects were likely to respond that the target and surround were consistent for all contrasts in undetected trials. We find no main effect of contrast ($p>.05$), meaning this bias persists despite an increased availability of feedforward information, which is reflective of the subjects' inability to detect the target region in these trials. In addition to evidence of a response bias, we find that subject sensitivity remains around chance level across all contrasts (with a non-significant main effect of contrast), meaning not only are subjects biased toward responding according to the surround, when this is controlled for, they are no longer able to distinguish the content of the target region.

Taken together, these results reveal differences in underlying behavioural measures between trials which were detected, versus those which were not. Subjects are still able to determine the content of the target region despite a small bias towards the surround in detected trials, whereas this ability is lost when they can no longer perceive the target region. Figure 2.11 summarises the SDT results for Undetected Trials.

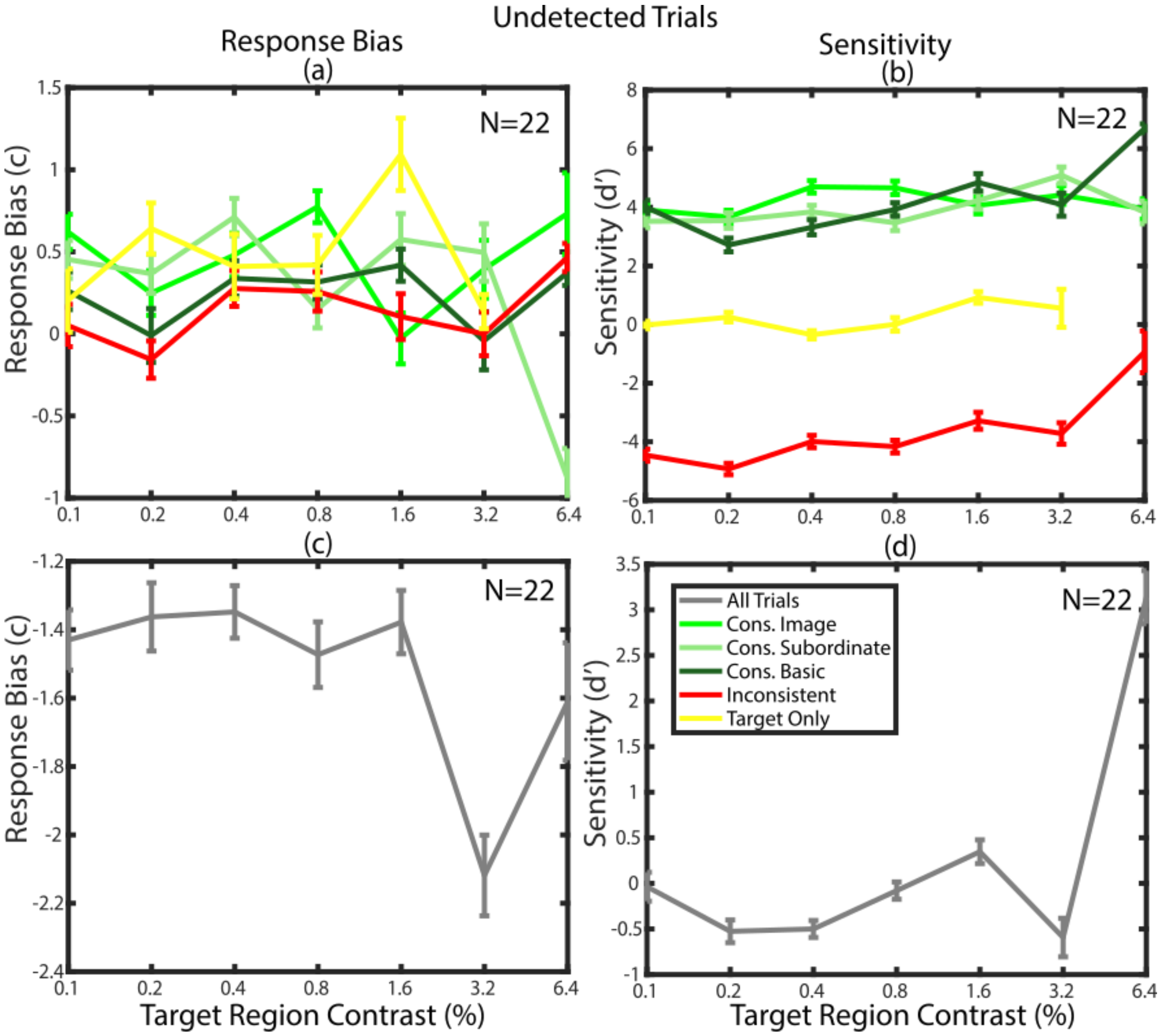


Figure 2.11 SDT results for undetected trials only. Response bias and sensitivity measures for The Consistency Effect Experiment in undetected trials only. The top row shows response bias (c) and sensitivity (d') in each consistency condition, at each contrast level. Here, hits, misses, false alarms and correct rejections are calculated using the content of the target region and the subject response. The bottom row shows overall response bias (left) and sensitivity (right) when hits, misses, false alarms and correct rejections are determined using the consistency between the surround and target regions. Error bars represent Standard Error.

2.4.2 The consistency effect persists around threshold contrast-level

2.4.2.1 Detection Frequency

The Threshold Contrast Experiment tests the effects of consistency around the threshold contrast level. We find a significant main effect of consistency in the detection task (F(1.251, 11.262)=26.071, p<.001) with a large effect size η_p^2 =.743. In the previous two experiments, we do not find any influence of consistency on detection frequency; however, here we see a higher detection rate in the consistency condition with no information in the surround (Target Only). Post-hoc pairwise comparisons using Bonferroni Correction reveal this difference is significant ($p<.005$), with the Target Only condition leading to significantly higher (Mean=71.713, SE=4.714) detection rates than Consistent trials (Mean=47.215, SE=4.009) and Inconsistent trials (Mean=49.787, SE=4.552), $p<.005$, both of which do not differ significantly from each other ($p>.05$). This detection enhancement effect for trials in which there is no information in the surround is not observed around the threshold in either of the previous experiments.

There is also a significant effect of contrast level ($F(1.871, 16.841)=123.516$, $p<.001$), $\eta_p^2=.932$, supporting our previous findings that detection frequency increases with contrast. Post-hoc pairwise comparisons reveal these differences exist between all contrast levels ($ps<.05$) except the lowest three contrasts and the highest two (i.e., those which are physically most similar to each other). We again observe a significant interaction between consistency x contrast ($F(2.482, 22.334)=3.829$, $p=.030$) $\eta_p^2=.298$. Figure 2.12 (left) indicates that as contrast increases, the effect of consistency decreases, and the detection frequency becomes more similar across consistency conditions when subjects can better perceive the low-contrast information.

2.4.2.2 Recognition Accuracy

Results of the recognition task in The Threshold Contrast Experiment Figure 2.12 (right) reveal the same response pattern observed in the previous experiments, with Consistent surround information resulting in the highest performance accuracy (Mean=81.274%, SE=3.893%), followed by the Target Only condition (Mean=67.426%, SE=4.878%) and finally the Inconsistent condition (Mean53.764%, SE=9.750%). Post-hoc pairwise comparisons using Bonferroni Correction reveal that these differences are not significant and the main effect of consistency on performance accuracy doesn't quite reach significance ($F(1.179, 9.436)=4.752$, $p=.055$).

There is a significant main effect of contrast ($F(1.774, 14.192)=11.606$, $p=.001$), $\eta_p^2=.592$. Post-hoc analysis reveals a significant difference only between the lowest and highest contrast level ($p<.05$). The interaction between consistency and contrast level is non-significant (F(4.138, 33.101)=0.66, p=.625), with differences between consistency conditions remaining stable across contrast levels.

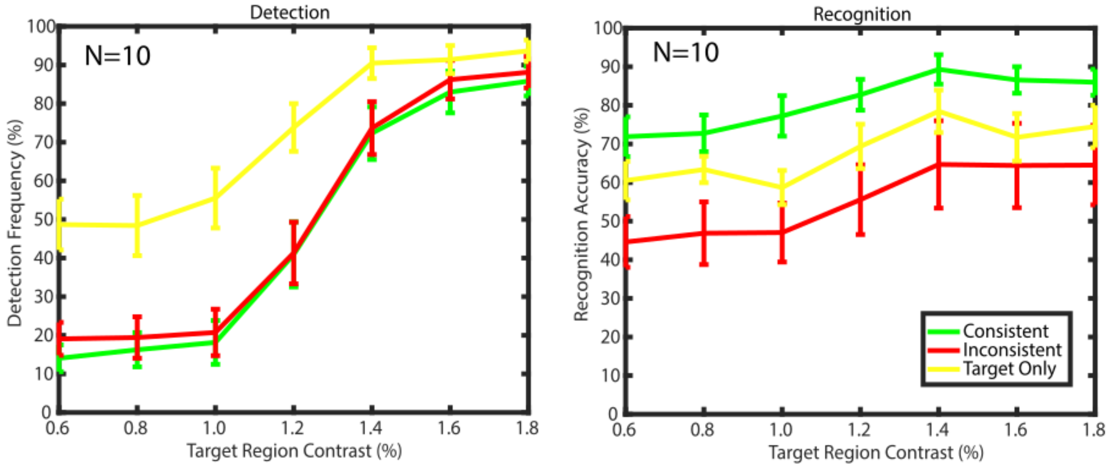


Figure 2.12 The Threshold Contrast Experiment detection and recognition results. Detection frequency (left) and recognition accuracy (right) in The Threshold Contrast Experiment. Results are depicted at for each consistency condition, at each contrast level. Error bars represent Standard Error.

2.4.2.3 Response Bias and Sensitivity

We will start with the application of Signal Detection Theory which takes only the target region content into account. Looking at response bias, we find no significant difference in criterion levels between the consistency conditions (F(1.167, 10.500)=0.251, p=.663). There is a significant effect of contrast (F(2.432, 21.886)=14.332, p<0.001), η_p^2 =.614, with response bias decreasing as contrast increases (revealed through significant post-hoc pairwise comparisons between the lower and higher contrast levels). In the lower contrast levels, subjects are more likely to respond that the target region contains a mountain in all of the conditions. This bias reduces to around chance as contrast levels increase. The interaction between consistency and contrast is insignificant (F(5.073, 45.655)=1.182, p=.333), meaning response bias did not vary between conditions at any contrast level.

Looking at sensitivity for each condition, we find a significant main effect of consistency in The Threshold Contrast Experiment (F(1.075, 9.671)=5.061, p.047), with an effect size of η_p^2 =.360. Sensitivity is improved when surround and target regions are consistent, relative to both no information in the surround and inconsistent information in the surround, however Bonferroni-corrected post-hoc comparisons reveal these differences to be non-significant. A significant effect of contrast was also found (F(2.284, 20.553)=5.098, p=.013), η_p^2 =.362, with sensitivity increasing with contrast level (again however, the post-hoc comparisons are non-significant). The interaction between consistency and contrast is non-significant (F(5.023, 45.206)=0.844, p=.526) meaning these consistency differences persist across all contrast levels.

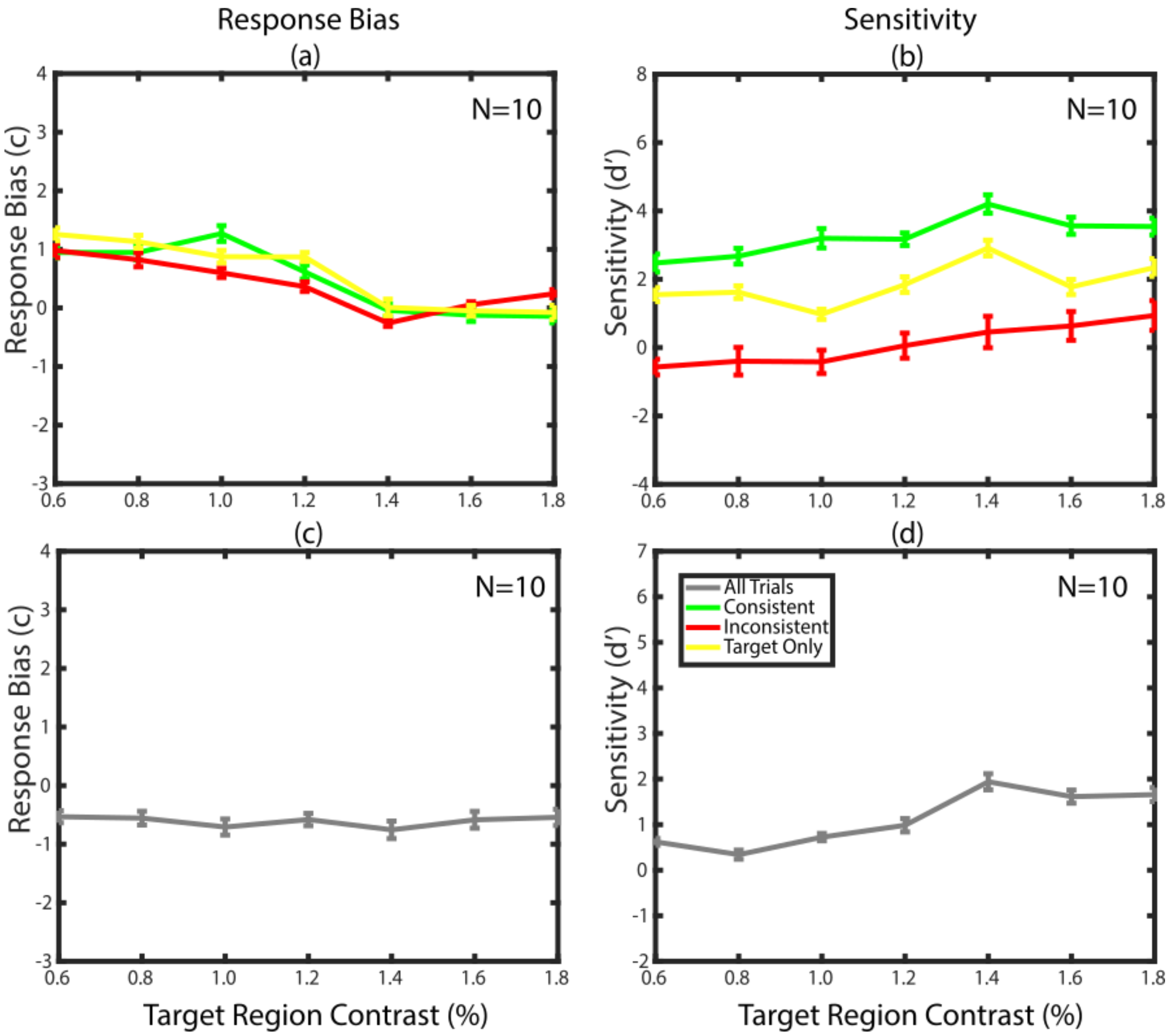


Figure 2.13 SDT results for the Threshold Contrast Experiment. Response bias (left) and sensitivity (right) results for the Threshold Contrast Experiment. The top row depicts these measures for each consistency condition, at each contrast level when hits, misses, false alarms and correct rejections are calculated using the content of the target region and subject response. The bottom row shows the results when the consistency between the surround and target regions is taken into account to calculate hits, misses, false alarms and correct rejections. Error bars represent Standard Error.

When considering the consistency between the surround and target regions as the basis of the signal detection analysis, we find that subjects are slightly more inclined to respond that the target region and surround are consistent across all contrast levels, therefore finding no significant effect of contrast on response bias (F(6, 54)=1.052, p=.403). This bias is only small, with criterion values remaining above -1 at all times. In addition to finding little evidence of a response bias, subject sensitivity is above chance level and increases with contrast, driving a significant effect of contrast on sensitivity (F(2.832, 24.487)=6.352, p=.003), η_p^2 =.414 (post-hoc comparisons, however, reveal these differences are largely insignificant (p>.05) expect between 0.8% contrast and both 1.6% and 1.8% (ps<.05). Figure 2.13 summarises the SDT-based results for the Threshold Contrast Experiment.

In sum, we again find little evidence of different contextual information having an influence on detection rates around the perceptual threshold. Instead, the absence of any information in the surround results in better detection of low-contrast information in the target region. We find a similar response pattern in terms of accuracy in the recognition task, with increased performance when surround information is consistent and decreased performance accuracy when it is inconsistent. This effect almost reaches significance. We find no evidence of a response bias difference between conditions and little evidence of a response bias towards reporting surround and target consistency overall. In a similar manner to accuracy, we find consistent information in the surround increases subject sensitivity and inconsistent information decreases it. Removing the potential bias towards reporting that the surround and target are consistent, we still find subject sensitivity to be above chance and this increases with contrast level.

2.4.3 The consistency effect extends beyond simple image features

2.4.3.1 Recognition Accuracy

The Beyond Simple Image Features Experiment consisted of only a recognition task and was performed on the threshold-contrast levels used in the previous. Results reveal a significant main effect of consistency (F(1.230, 33.298)=22.864,

$p<.001$), with a large effect size of $\eta_p^2 =.459$, reported here with Greenhouse-Geisser corrections. Reflective of our previous findings, we find consistency between surround and target increases recognition performance, with highest accuracies in the Consistent Subordinate (e.g., beach and another beach) condition (Mean=77.091%, SE=2.040%) and in the Consistent Basic (e.g., beach and mountain) condition (Mean=77.228%, SE=2.308%). The facilitatory effect does not differ significantly between these two consistent conditions ($p>.05$). The Inconsistent condition results in the lowest performance (Mean=57.354%, SE=6.636%), although here at around chance level rather than below chance as in previous experiments, and the Target Only accuracy falls between the Consistent and Inconsistent conditions (Mean=70.072%, SE=2.222%). Differences between all consistency levels are significant ($ps<.05$) except the two Consistent conditions, which both exert similar facilitatory effects on performance accuracy ($p>.05$).

We also find a significant effect of contrast in this experiment, with increased performance with increasing contrast levels ($F(2.282, 61.616)=40.831$, $p<.001$), $\eta_p^2 =.602$, with pairwise comparisons revealing a significant distinction between low versus high contrast levels ($ps<.05$). The interaction between consistency and contrast is non-significant ($F(7.849, 211.929)=1.796$, $p=.081$); meaning the consistency effect between consistency conditions remains stable across contrast levels. Performance accuracy is summarised below in Figure 2.14.

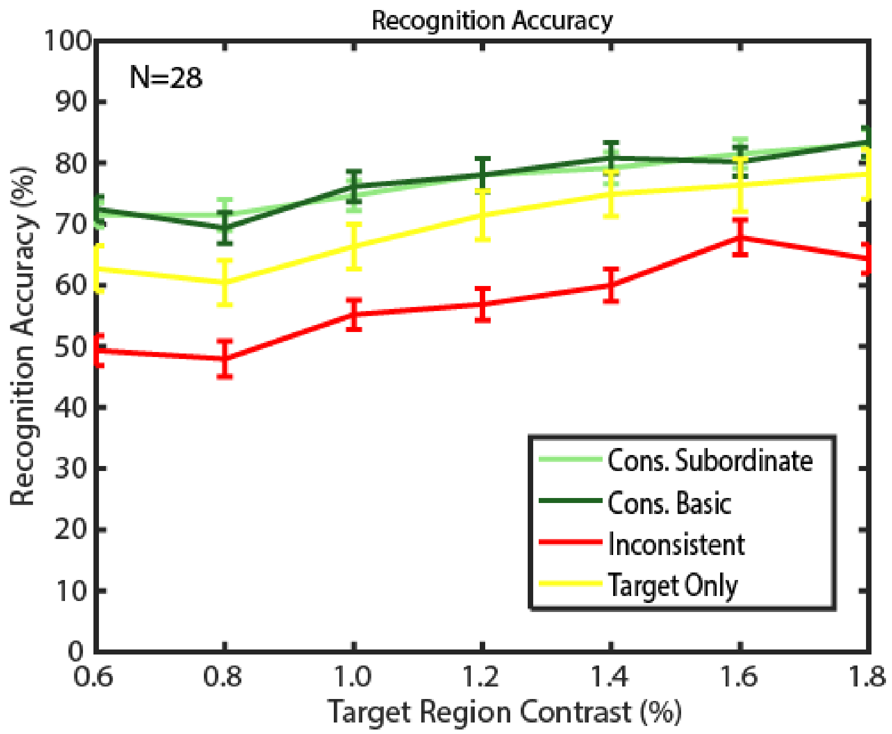


Figure 2.14 Recognition accuracy for The Beyond Simple Image Features Experiment. Subject performance is displayed as a function of contrast level for each consistency condition. Error bars represent Standard Error.

2.4.3.2 Response Bias and Sensitivity

Using the content of the target region to run the Signal Detection analysis Figure 2.15 (top), the effect of consistency on response bias almost reaches significance (reported using Greenhouse-Geisser corrections) - (F(2.337, 60.758)=2.989, p=.50), $\eta_p{}^2$ =.103. However, the overall span of response bias is low and the majority of differences between conditions are non-significant (p>.05, reported using Bonferroni corrected pairwise comparisons). Subjects are slightly more inclined to report that the target contained a beach similarly across all conditions. This response bias decreases significantly with contrast increase (F(3.884, 100.988)=3.575, p=.010), $\eta_p{}^2$ =.121 (although post-hoc comparisons reveal no significant differences between conditions (ps>.05), but does so uniformly across all consistency conditions, resulting in a non-significant interaction between consistency and contrast. When looking at sensitivity, we find a significant main effect of consistency on sensitivity (F(1.527, 39.703)=21.858, p<.001) with a large effect size of $\eta_p{}^2$ =.457, with subject sensitivity being higher in both of the Consistent conditions (Consistent Subordinate: Mean=2.240, SE=.262; Consistent Basic: Mean=2.293, SE=.266) and lower (around chance) in the Inconsistent condition (Mean=.689, SE=.249). Post-hoc Bonferroni-corrected comparisons reveal significant differences between all conditions (ps<.05) except the two Consistent conditions (p>.05).

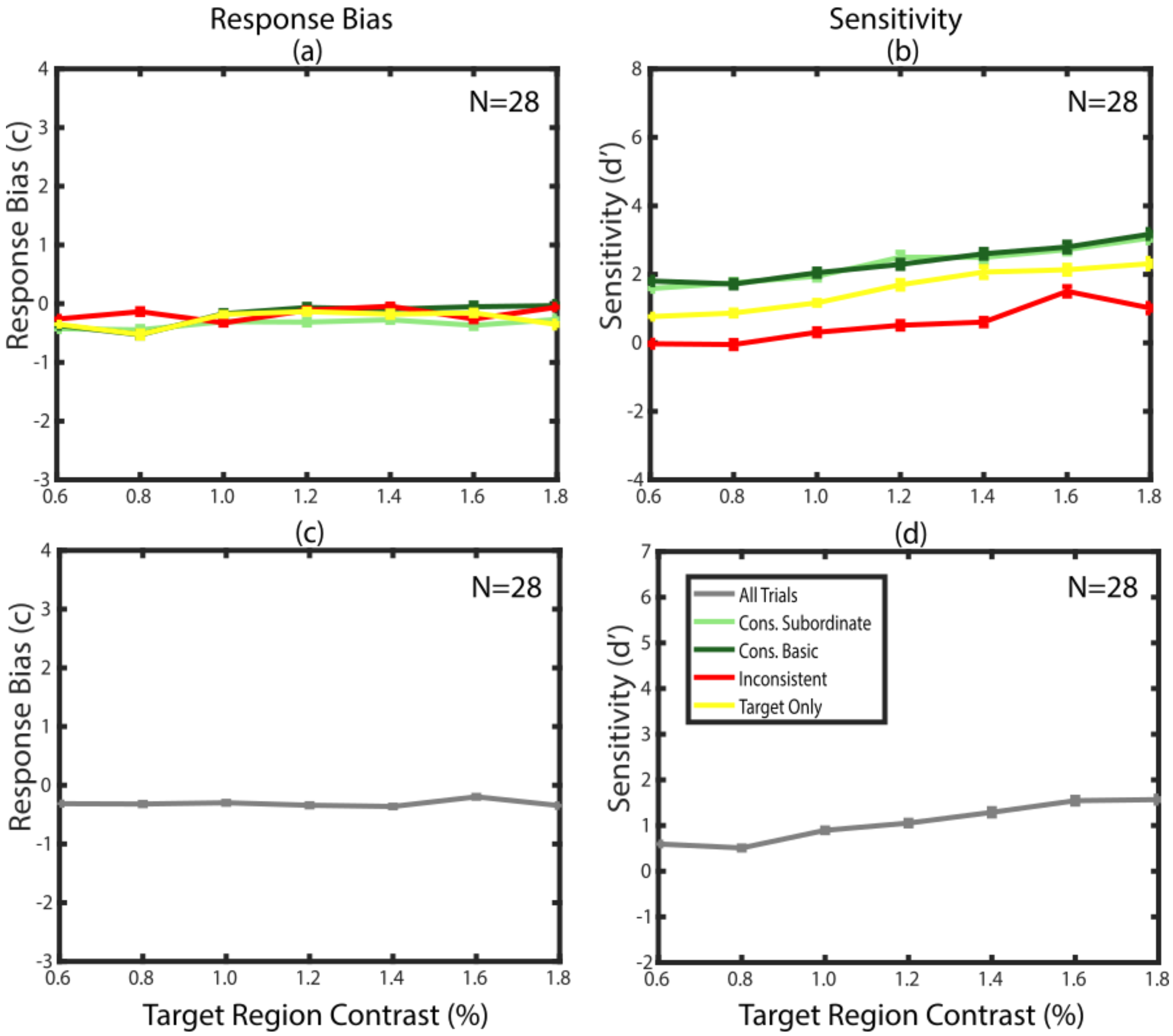


Figure 2.15 SDT results for the Beyond Simple Image Features Experiment. Response bias (left) and sensitivity (right) results for The Beyond Simple Image Features Experiment. The top row depicts bias and sensitivity within each consistency condition when hits, misses, false alarms and correct rejections are calculated based on the subject response and the content of the target region. The bottom row shows bias and sensitivity when the consistency between the surround and target regions is used to determine d' and criterion values. Error bars represent Standard Error.

When taking the consistency between the surround and target region into account within the Signal Detection framework Figure 2.15 (bottom), overall response bias is low however, we find a slightly significant effect of contrast (F(3.850, 103.956)=2.545, p=.046), η_p^2 =.086, with subjects being less likely to respond that the target and surround are consistent when the contrast is higher (i.e. the target region is more visible). Post-hoc comparisons of the contrast levels are largely insignificant except for 1.6% contrast which differs significantly from the contrast levels surrounding it (ps<.05). Sensitivity is low when contrast levels are low, but we find a highly significant effect of contrast on sensitivity (F(3.202, 86.445)=31.027, p<.001), η_p^2 =.535, with subjects having increased sensitivity with increased contrast and post-hoc pairwise comparisons revealing significant differences between the lower and higher contrast levels (ps<.05).

Overall, we replicate the findings of the Consistency Effect Experiment whereby consistency improves recognition accuracy despite simple image features like line continuation not being used to determine the consistency between the two regions. This enhancement in the consistent conditions persists in the measure of sensitivity, with higher sensitivity when the surround and target are consistent. We find an above chance level of sensitivity when response bias for reporting that the target is the same as the surround is taken into account. This increases when the target region becomes more visible, and the response bias decreases with increased contrast.

2.4.3.3 Response Bias Experiment

A General Linear Mixed Model was used to assess performance accuracy as a function of contrast. The GLMMs fitted logistic psychometric functions to determine the probability of participants correctly identifying the target stimulus in each condition. In each condition, performance was compared at the 25%, 50% and 75% points (Q.25, Q.50 and Q.75, respectively) on each psychometric curve, (note that these are percentages above chance performance).

Results reveal that overall performance increases with contrast level, as would be expected. Furthermore, context plays a more salient role in identification accuracy at the lower contrast levels, with a performance increase being observed in the Same Target condition when the contrast levels are low (at the Q.25 level), compared to all other conditions. As contrast increases, this Same Target condition advantage decreases, implying that the role of context becomes less influential when the feedforward information is more visible, with performance being similar across all conditions at the Q.25 and Q.5 levels. Consistency in general, did not influence performance.

In this paradigm, contextual information only facilitates performance in the condition where line continuation is possible. Contextual consistency alone (with the absence of line continuation) is insufficient in improving task performance. The line continuation advantage does not persist in the same non-target condition, implying an element of task relevancy. Further, category consistency did not generally improve performance. Figure 2.16 summarises results for the Response Bias Experiment.

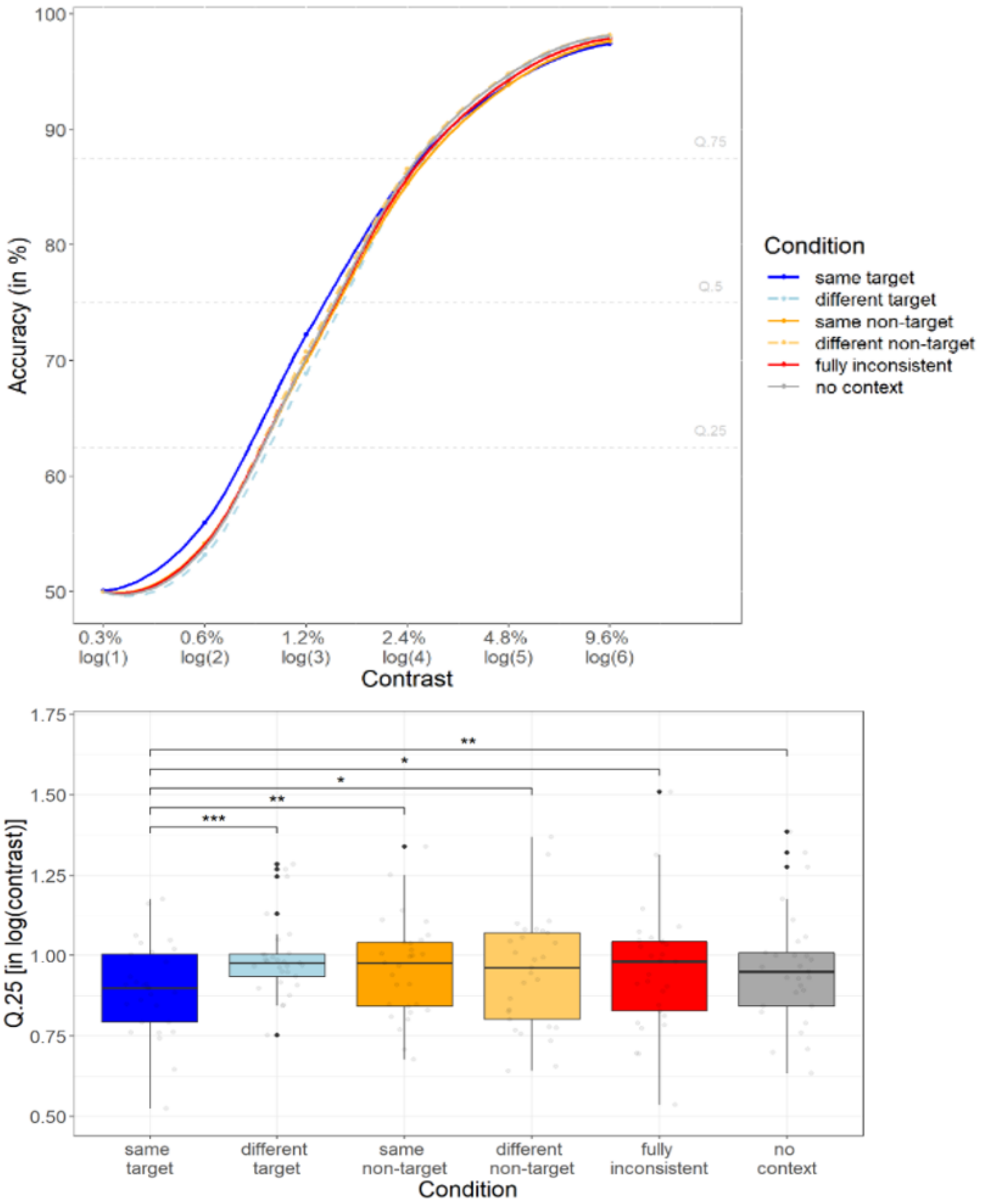


Figure 2.16 Response Bias Experiment results – a summary of model fit for the MSc Experiment conducted within the lab. The top figure depicts the psychometric curves fitted by the GLMM for each condition. The curves predict accuracy values as a function of contrast level. The bottom figure compares condition performance at the Q.25 level in which the same target condition provides an advantage. Figures obtained with permission from the MSc student who produced them as part of their report. N=29.

2.4.4 Altering spatial frequencies

2.4.4.1 Spatial Frequency manipulated in the target region

In this experiment (The Spatial Frequency Target Experiment), the image in the surround was always shown at full contrast (except the Target Only condition, in which no scene was shown in the surround). Spatial frequency of the target region was manipulated, such that it was shown at low contrast but also at either high or low spatial frequency. Again, subjects only performed a recognition task in this experiment.

We performed a three-way repeated measures ANOVA to investigate the individual influences of the independent variables of spatial frequency (low versus high), consistency and contrast level, as well as any interactions between these predictors. Results are summarised below in Figure 2.17. When considering performance accuracy, we found a significant and large effect of spatial frequency (F(1.000, 9.000)=12.170, p=.007) η_p^2 =.575. Post hoc comparisons using Bonferroni corrections reveal that low spatial frequency in the target region resulted in significantly (p<.05) higher performance accuracy (M=63.305%, SE=4.069%) than a high spatial frequency target region (M=54.320%, SE=2.187%), regardless of consistency or contrast level. Generally, we also observe an overall lower performance accuracy in this experiment than any of our previous experiments, suggesting limiting the spatial frequency of the low-contrast information increases task difficulty.

We find no significant main effect of consistency (F(2, 18)=.531, p=.056) or contrast (F(6, 54)=2.247, p=.052), and therefore do not replicate the findings of the previous experiments. Unlike our previous results, consistency between surround and target regions does not influence the subjects’ ability to identify the low-contrast information. Performance does improve slightly with increasing contrast-level; however, this does not reach significance, indicating limiting spatial frequency overrides the beneficial effects of a more visible stimulus. We also find no significant interactions between any of the predictors.

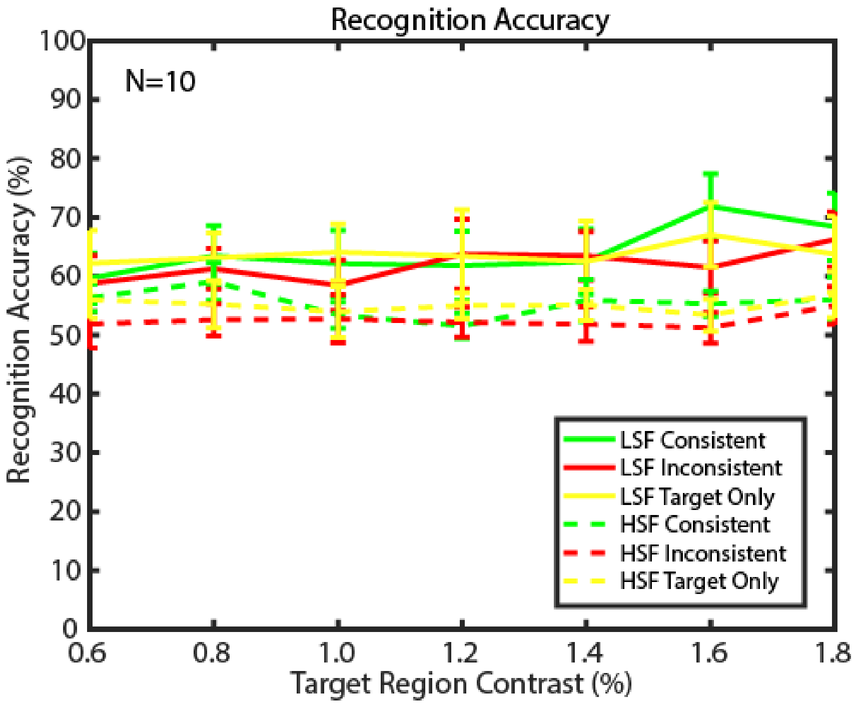


Figure 2.17 Performance accuracy in the Spatial Frequency Target Experiment. Spatial frequencies limited (to low versus high) in the target region. Performance accuracy in the recognition task is depicted for each consistency condition and each spatial frequency type, at each contrast level. Error bars represent Standard Error.

Analysis of the data using signal detection theory reveals some key influences of spatial frequency which can be observed in Figure 2.18. When considering response bias in terms of the content of the target region only (whereby a negative response bias indicates that subjects are more likely to respond that the target contains a beach and a positive bias indicates subjects are more likely to respond with mountain), we find a significant effect of spatial frequency (F(1.000, 9.000)=14.787, p=.004), η_p^2 =.622 which is confirmed with post-hoc analysis which is also significant (p<.005). Criterion values for both spatial frequency types are largely positive, meaning subjects are generally more likely to consider the target region as a mountain over a beach. This bias is significantly more pronounced when the target region is limited to high (Mean=2.333, SE=.423) rather than low (Mean=1.315, SE=.536) spatial frequencies. We do not find a significant main effect of consistency condition (F(1.096, 9.868)=.835, p=.394), or contrast (F(2.517, 22.656)=1.605, p=.220) meaning the consistency between the surround and target region or the contrast of the target region itself does not influence subjects' criteria for determining their response.

We do however find a significant interaction between spatial frequency and consistency condition (F(1.138, 10.242)=15.057, p=.002), η_p^2 =.626, with almost mirrored criterion patterns between conditions within the two spatial frequency types. This is especially evident in the condition in which there is no information in the surround (Target Only condition – yellow lines) whereby high spatial frequency information in the target region results in subjects being more likely to respond with mountain and low spatial frequency target information means subjects are more likely to respond with beach. This interaction implies overall that the consistency between the target and surround modulates the impact of spatial frequency limitations in subjects' ability to identify the target region.

We also find a significant interaction between spatial frequency and contrast (F(2.320, 20.876)=4.811, p=.016), η_p^2 =.348, with response bias discrepancies between the two conditions increasing with contrast level. That is, as the content of the target region becomes more visible, subjects are more likely to respond differently according to the type of spatial frequency information provided in the target region.

We do not find a significant interaction between consistency and contrast level (F(29.066, 4.218)=1.059, p=.385) or between all three factors (SF x Consistency x Contrast), (F(4.218, 37.858 = 1.955, p=.118).

In terms of sensitivity as a measure of performance, we find a significant effect of spatial frequency on d' values (F(1.000, 9.000)=11.442, p=.008), η_p^2 =.560, with low spatial frequency information in the target region resulting in increased sensitivity (Mean=1.464, SE=.463) relative to high spatial frequency information (Mean=.546, SE=.269) – supported by the Bonferroni-corrected pairwise comparisons (p<.05). We do not find a significant main effect of consistency (F(2, 18)=.961, p=.401) or contrast (F(6, 54)=1.291, p=.277). We find no significant interactions between any of the factors (ps>.05). These results imply that the type of spatial frequency information in the target region influences subjects' ability to determine its content, but the contextual information in the surround and the relative visibility of the target region in terms of contrast do not. Therefore, sensitivity is driven primarily by what spatial frequency information is available and not other low-level image features such as contrast or higher-level influences such as context.

When we take the consistency between the surround and target region into account within the Signal Detection Theory parameters, a negative response bias indicates that subjects are more likely to respond that the target region is consistent with the surround. Response bias in this instance is very small and around chance level for both spatial frequency types. We do not find a significant difference in subject criterion values between having low versus high spatial frequency information in the target region (F(1.000, 9.000)=.629, p=.448). We also find no significant main effect of contrast (F(1.627, 14.643)=1.866, p=.193) nor a significant interaction between spatial frequency and contrast level (F(1.272, 11.449)=1.065, p=.343). Therefore, there is no evidence of a response bias toward responding that the surround and target region are consistent in this experiment.

Using sensitivity to measure performance, we again find a significant difference between the two spatial frequency types with low spatial frequency resulting in higher sensitivity across subjects (Mean=.882, SE=.362) than high spatial frequency information (Mean=.211, SE=.107), (F(1.000, 9.000)=6.565, p=.031), η_p^2 =.422. This means subjects are better at determining the content of the target region when low spatial frequency information is preserved (supported by significant (p<.05) post-hoc comparisons). Sensitivity does increase with contrast level in both spatial frequencies however this effect does not reach significance (F(6, 54)=1.888, p=.100) and we do not find a significant interaction between spatial frequency and contrast (F(2.298, 20.680)=1.149, p=.343).

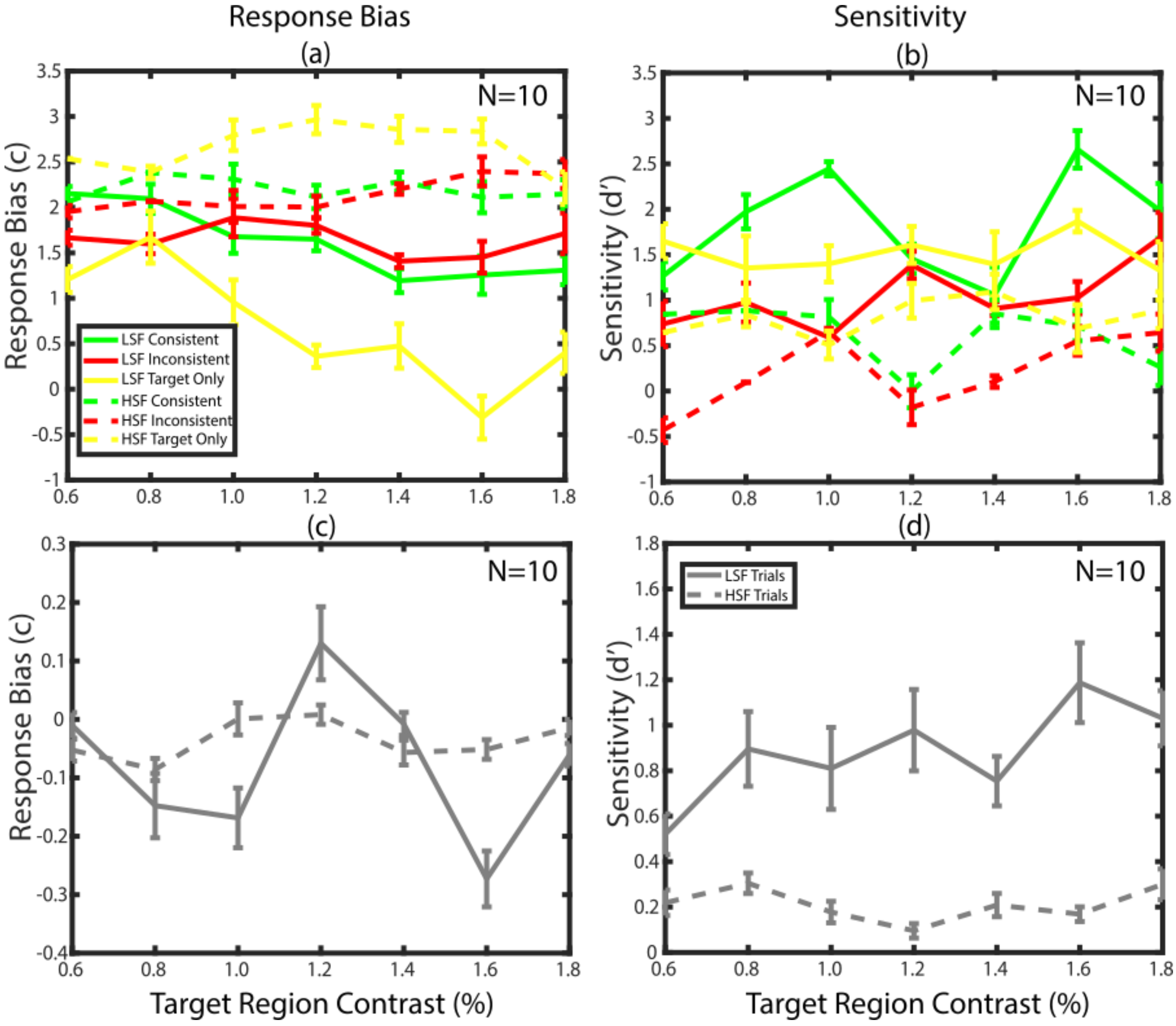


Figure 2.18 SDT results for the Spatial Frequency Target Experiment. Response bias (left) and sensitivity (right) in the Spatial Frequency Target Experiment in which spatial frequencies were limited in the target region. The top row depicts criterion and d' within each consistency condition, for each spatial frequency as a function of contrast level. With hits, misses, false alarms and correct rejections calculated using the content of the target region and subject response. The bottom row shows the bias and sensitivity values as a function of contrast level for high and low spatial frequency data. Here, hits, misses, false alarms and correct rejections are calculated using the consistency between the surround and target regions. Error bars represent Standard Error.

In sum, we find that low spatial frequency information in the target region significantly improves both accuracy and sensitivity in being able to determine the scene depicted. This effect seems to override the beneficial influence of consistent information in the surround; however, this could be attributed to the increased task difficulty as a result of further degradation of the target region. Although so far, we have generally found more of a dependence upon the surround when the target region is most degraded, manipulation of the spatial frequency information available in the target region appears to diminish this effect.

2.4.4.2 Spatial Frequency manipulated in the surround region

The Spatial Frequency Surround Experiment manipulated spatial frequencies in the surround. The target region was shown at low contrast but spatial frequencies were not limited. Instead, the surround was shown at either low or high spatial frequency (but at full contrast, as in the previous experiments).

Performance accuracy results (Figure 2.19) reveal a significant main effect for spatial frequency on response accuracy ($F(1.000, 9.000)=6.156$, $p=.035$), $\eta_p^2 =.406$. In contrast to The Spatial Frequency Target Experiment, performance accuracy is significantly higher when spatial frequency is limited to higher ($M=73.030\%$, $SE=4.966\%$) rather than low ($M=68.643\%$, $SE=4.469\%$) levels (supported by significant ($p<.05$) post-hoc results). This is particularly evident in the Consistent conditions, however the interaction between spatial frequency and consistency is insignificant ($p>.05$). Looking at consistency itself, post-hoc Bonferroni-corrected pairwise comparisons reveal higher accuracy in the Consistent condition ($M=80.128\%$, $SE=3.661\%$) than the Inconsistent condition ($M=61.545\%$, $SE=8.545\%$), in line with previous findings. However, this effect fails to reach significance ($F(1.000, 9.000)=3.983$, $p=.077$).

We do observe a significant effect of contrast ($F(1.549, 13.941)=11.039$, $p<.005$), $\eta_p^2 =.551$, with performance accuracy increasing with contrast level (although Bonferroni-corrected pairwise comparisons reveal these differences to be mainly insignificant ($ps>.05$). We also find a significant interaction between spatial frequency and contrast ($F(6, 54)=3.667$, $p=.004$), $\eta_p^2 =.289$, with differences in spatial frequency effects reducing as contrast of the target region increases.

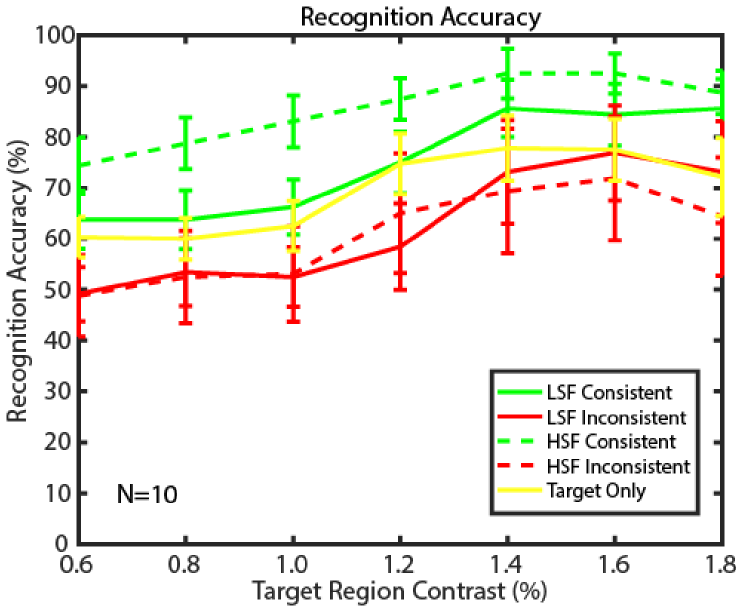


Figure 2.19 Performance accuracy in The Spatial Frequency Surround Experiment. Performance in the recognition task in The Spatial Frequency Surround Experiment in which spatial frequency information is limited in the surround and the target region is shown at low contrast. Accuracy is displayed for each consistency condition for both types of spatial frequency information, as a function of contrast level. Error bars represent Standard Error.

Signal Detection Analysis taking only the target region content into account can measure differences in response bias and sensitivity between consistency conditions. As noted, a negative response bias here indicates that subjects are more likely to respond that the target region is a beach. We find a significant main effect of spatial frequency on response bias (F(1.000, 9.000)= 5.690, p=.041), η_p^2 =.387, with subjects significantly more likely to respond that the target region is a mountain when the surround is limited to low spatial frequencies (or absent as in the Target Only condition), (Mean=1.601, SE=.376), than when the surround contains only high spatial frequencies (Mean=.273, SE=.328), as confirmed by post-hoc comparisons (p<.05). This bias significantly reduces as contrast increases (F(1.860, 16.739)=12.164, p=.001), η_p^2 =.575, and is around chance at the highest contrast levels. Post-hoc comparisons reveal the higher contrast levels to mainly differ significantly from the other contrasts. We also find a significant interaction between spatial frequency and contrast (F(6, 54)=2.483, p=.034), η_p^2 =.216, with the distinction between the spatial frequency types reducing as contrast increases. The bias does not differ significantly between consistency conditions (F(1.000, 9.000)=.962, p=.352) and we find no significant interactions between consistency and spatial frequency or consistency and contrast level (ps>.05). We do find a significant interaction between all three factors (F(6, 54)=2.282, p=.049), η_p^2 =.202. Differences between the spatial frequencies decrease as the target region becomes more visible and at the higher contrast level a consistent surround and target region mean that subjects are more likely to respond with beach over mountain and vice versa for the Inconsistent condition. These differences are slight however, and overall response bias at the higher contrast levels is near chance level. The interaction itself only just reaches significance.

In terms of sensitivity, we only find a significant effect of contrast (F(1.995, 17.958)=8.862, p=.002), η_p^2 =.496, with sensitivity increasing with contrast. Mean sensitivity is higher when the surround and target regions are consistent (4.043, SE=.399) compared to when they are inconsistent (Mean=1.388, SE=1.113), however, this effect does not reach significance (F(1.000, 9.000)=4.604, p=.074). There is no significant main effect of spatial frequency, nor any significant interactions between variables. This implies that when contextual information is limited in terms of its spatial frequency and response bias is taken into account, we do not find a significant influence of consistency on the ability to identify the low-contrast target region.

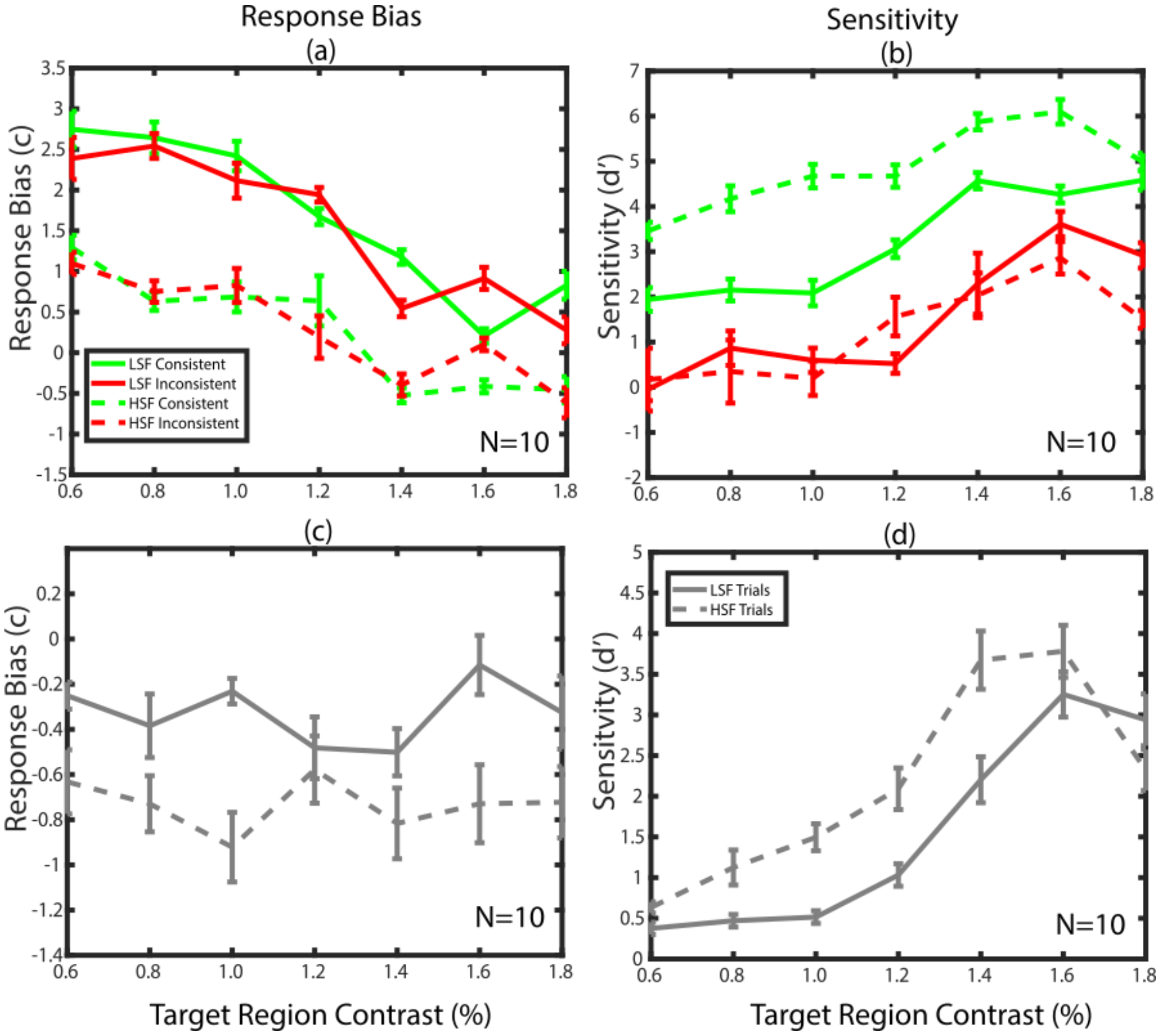


Figure 2.20 SDT results for the Spatial Frequency Surround Experiment. Response bias and sensitivity in the Spatial Frequency Surround Experiment in which spatial frequency information in the surround is limited to either low or high spatial frequencies and the target region is displayed at low contrast but with spatial frequency information not limited. The top row depicts bias and sensitivity as a function of contrast level within each consistency condition and for both spatial frequency groups. Hits, misses, false alarms and correct rejections are calculated depending on the content of the target region and the subject response. The bottom row depicts bias and sensitivity for each spatial frequency category (low versus high), as a function of contrast level, with consistency between the surround and target region being used to determine these measures. Error bars represent Standard Error.

When consistency between the surround and target region is used to calculate signal detection measures, we find no significant main effects of response bias for either spatial frequency (F(1.000, 9.000)=1.567, p=.242) or contrast (F(3.665, 32.981)=.361, p=.819). We also do not find a significant interaction between these factors (F(6, 54)=.602, p=.727). Response bias overall is negative, meaning subjects are more likely to indicate that the target is consistent with the surround, however, this bias is only slight with criterion values remaining above -1.

With sensitivity, we find a significant main effect of spatial frequency (F(1.000, 9.000)=7.578, p=.022), η_p^2 =.457, with higher spatial frequencies in the surround eliciting higher sensitivity in being able to determine the content of the target region (Mean=2.161, SE=.430) relative to low spatial frequency information in the surround (Mean=1.541, SE=.395), supported by post-hoc pairwise comparisons (p<.05). We also find a significant main effect of contrast (F(2.125, 19.126)=8.475, p=.002), η_p^2 =.485, with sensitivity increasing with contrast level (although post-hoc comparisons are mainly insignificant except 1.6% contrast differing significantly from 0.6% and 0.8%). The interaction between spatial frequency and contrast is non-significant (F(6, 54)=1.706, p=.137), meaning the distinction between the two frequency types remains stable across contrast levels. SDT-based results for The Spatial Frequency Surround Experiment are summarised in Figure 2.20.

To sum, in contrast to the Spatial Frequency Target Experiment, in the Surround Experiment we find that the high spatial frequency information results in better performance across both accuracy and sensitivity measures. That is, high spatial frequency contextual information allows subjects to more accurately determine the content of degraded feedforward information.

2.5 Discussion

2.5.1 Summary of Results

2.5.1.1 Consistent context enhances recognition (but not detection) of low-contrast information

Results from the Consistency Effect Experiment and the Simplified Experiment indicate a differential influence of context upon detection and recognition of degraded visual input. The consistency of the high-contrast contextual information does not influence subjects' ability to detect low-contrast information in the target region. That is, the type of information in the surround did not influence subjects' ability to detect information in the target region as detection rates were similar across consistency conditions. This is not in line with previous findings which have observed a modulatory effect of congruency on detectability (Caddigan et al., 2017; Hollingworth & Henderson, 1998). Caddigan and colleagues demonstrated that the degree to which an image exemplifies its category influenced how easily it was detected. Using intact or scrambles images, subjects detected intact images better regardless of whether categorisation was a task factor or not. Perhaps this effect extends to how well intact images represent their categories and could highlight a limitation with our specific images. Their detection task was also based on rapid presentation of the (full contrast) image, meaning the amount of signal was greater than in our experiment in which the signal was very low in the low-contrast target region. Hollingworth & Henderson found a perceptual advantage for semantically consistent objects and scenes. Our stimuli consistent of images of scenes without objects which may incorporate a different level of visual processing. (Note that Hollingworth & Henderson's effect could also be explained by a response bias and therefore may not truly represent a perceptual advantage in sensitivity for semantically consistent objects). We find that contrast level does have an influence on detection, as one would expect, as the low-contrast information becomes more visible (with increased contrast), subjects' detection frequency increases, reflective of a general contrast effect (Grill-Spector & Malach, 2004a).

Congruency of available information in the surround does have a significant effect on the ability to recognise the content of the low-contrast ('occluded') region. Consistent information in the context significantly improves subjects' ability to identify the low-contrast scene presented in the target region, relative to having no contextual information available. Conversely, having inconsistent information in the context significantly decreases subjects' ability to identify low-contrast information relative to the condition in which no context is present. This implies that when feedforward information is difficult to perceive, contextual information is used to provide predictions about the degraded input. When these predictions are relevant, the ability to determine the low-contrast information increases. On the other hand, when the predictions are irrelevant (based on inconsistent contextual information), the ability to identify the degraded feedforward information is hindered. Both of these findings suggest contextual modulation takes place to some degree when feedforward input it difficult to interpret. Increased contrast also improves performance in the recognition task, and this enhancement is stable across consistency conditions. This effect is more pronounced in The Consistency Effect Experiment than The Simplified Experiment which does not reach significance. This could be due to low power in The Simplified Experiment due to lower subject numbers or could reflect a fundamental difference in the design of the experiments.

The within-subjects design of the Consistency Effect Experiment allowed us to separate trials in which subjects detected a scene present versus those in which they did not. This enabled recognition performance to be assessed in stimuli which was either detected or not. Splitting the trials by whether or not they were detected does not influence the apparent effect of consistency upon recognition performance. We still observe a clear recognition advantage in trials with a consistent surround, regardless of whether or not the low-contrast scene was detected. In undetected trials however, visibility of the target region has less of an influence, with contrast level not exerting a significant effect. This would be expected as although visibility of the stimulus varies, this is irrelevant in trials in which subjects did not perceive a stimulus presence. In detected trials, recognition increases with contrast, as we observe when all trials are included.

The consistency effect therefore persists even in trials that were not detected by subjects, suggesting that this effect could extend into subconscious perception. However, SDT analysis reveals that this consistency effect can be explained by a response bias toward the surround in the undetected trials. Importantly, inferring consciousness from subject responses to a detection task may not be entirely accurate due to our relatively long presentation times and the assumption that subjects reported their visual experience reliably meaning conscious and subconscious processing cannot be accurately disentangled within the parameters of this experiment. To do so we would need to explore different ways to prevent conscious perception of the stimuli on some trials through different means such as rapid presentation or perceptual masking. We also find a strong response bias of subjects reporting that the surround and target regions are consistent within these trials which is coupled with sensitivity which is around chance level. This effect is not observed in detected trials which show a smaller response bias which decreases with contrast alongside increased sensitivity, which is more reflective of the SDT results when all trials are combined. This distinction between detected and undetected trials highlights the importance of considering the underlying components of behaviour. On face value, accuracy scores suggest that a consistency effect exists even in undetected trials, but SDT reveals this to be due to a response bias in which subjects are reporting the information in the surround.

During the recognition task, subjects are presented with a visual scene which for the most part is high contrast, but the task is aimed at information which is shown at very low contrast. It is therefore possible that subjects are (either implicitly or explicitly) overlooking the information in the target and attending only to the information in the surround, indicating the possibility of a response bias. Should this be the case, we would observe a similar response pattern, in that accuracy for the Consistent condition would be high and accuracy for the Inconsistent condition would be low. Performance should be around chance in the Target Only condition. This is the response pattern we find continuously throughout our experiments. A very sensitive observer would have high performance accuracy, as would a very biased observer. Thus, accuracy alone cannot provide a true representation of underlying behaviour.

In order to ascertain the extent to which a response bias contributed toward our consistency effect, we applied Signal Detection Theory. Signal Detection Theory uses participant hit ratio and false alarm ratio to compute two measures of behavioural response: response bias and sensitivity. Response bias refers to how likely participants are to respond in a certain manner and sensitivity refers to how well subjects are able to distinguish one stimulus type from another, whilst taking any bias into account (Swets et al., 1978). Evidence of a response bias has been found in previous studies investigating congruency effects (Bar, 2004; Henderson & Hollingworth, 1999; Hollingworth & Henderson, 1998) and it is thought that predictions bias perception in situations of uncertainty (Panichello et al., 2013), which our degraded stimulus lends itself to. Therefore, quantifying the contribution of a bias to our findings using SDT helps us understand the underlying mechanisms driving behaviour.

We used two applications of Signal Detection Theory within our analysis. Firstly, we split the data into response categories according to information in the target region (see Table 2.1 for details). This allowed us to measure whether subjects were biased towards reporting a certain scene in the target region when information in the surround was either consistent or inconsistent. It also allows a measure of sensitivity in being able to distinguish natural versus manmade (or beach versus mountain) images in the target region and whether the consistency differences we observe when looking at accuracy results persists within sensitivity.

In The Consistency Effect Experiment, we find little evidence of a response bias in any consistency conditions, with criterion values remaining close to zero across all contrast levels. Response bias results of The Simplified Experiment are more distinct and worth mentioning. The criterion values in this experiment suggest that when contrast levels are low, subjects are more likely to respond that the target is a mountain and as contrast levels increase (i.e., the target becomes more visible), subjects are more likely to respond that the target is a beach. This bias is reflected in the accuracy ratings where we see a slight decrease in accuracy at the higher contrast levels. Without any logical explanation as to why a more visible stimulus results in poorer performance that is biased in such a way that subjects respond inaccurately at the higher contrast levels, it is possible that the stimulus images used may be driving this observation. The Consistency Effect Experiment used several stimuli from different basic and sub-ordinate level categories. The Simplified Experiment used two images to generate the stimuli: one of a beach and one of a mountain. It is possible that at higher contrast levels, the beach and mountain stimulus images look very similar and are interpreted as a beach by the observer. The target region is positioned in the bottom right quadrant, an area of a mountain scene that would not give any obvious clues as to the identity of the scene, unlike for example, the upper quadrants of a mountain scene which would contain depictions of mountain peaks. It is therefore possible that the grassy side of the mountain can be easily mistaken for the sand of a beach when the target region becomes more visible.

Despite this potential bias, in The Consistency Effect Experiment we find differing sensitivity values (d’) between different conditions, with consistent information in the surround resulting in greater sensitivity in being able to determine the content of the target region and inconsistent information in the surround decreasing sensitivity to below chance. Sensitivity was around chance in the condition in which no contextual information was present, reinforcing the influence of contextual information being present. These distinctions, however, reduce with increased contrast suggesting the higher visibility of the target region reduces the impact of the contextual information surrounding it. Overall sensitivity across conditions increases with contrast meaning subjects were better able to distinguish the low-contrast information when it was easier to perceive. This increased sensitivity with contrast is replicated in The Simplified Experiment, but the consistency effect is not. We do see a slight sensitivity advantage in the consistent condition, but it does not reach significance and generally sensitivity across conditions is around chance until the target region is sufficiently visible.

The second application of SDT the consistency between the surround and target regions into account. So far, our results have only been able to show whether subjects are generally biased toward responding in a certain manner within each consistency condition, without factoring in the content of the surround itself. Ultimately, we wanted to know whether subjects were biased towards responding according to the surround, or, in other words, thinking the surround and target regions were consistent. We therefore recalculated hit and false alarm rates based on consistency between the surround and target regions (see Table 2.2 for details). This analysis, however, cannot report response bias and sensitivity within each condition, but can calculate these parameters overall to give an overall indication of subject responses when consistency between the surround and target is taken into account.

In terms of response bias, a more negative bias implies that subjects are more likely to (correctly or incorrectly) respond that the surround and target are consistent. This gives us a good indication as to whether subjects are responding according to the surround. In The Consistency Effect Experiment a slight bias toward responding in this way can be observed across contrast levels but decreases with increased contrast. Sensitivity is around chance at the lower contrast levels but increases considerably around the perceptual threshold. The higher contrast levels however are where we observe less of an effect of consistency. In The Simplified Experiment, the bias remains around chance across contrast levels but does increase slightly with increased contrast which again could reflect the stimulus images used. Sensitivity increases with contrast in The Simplified Experiment, but we do not observe a consistency effect in this Experiment as revealed by the initial SDT analysis.

Thus, our results point toward a potential response bias toward the surround at the lower contrast levels. This is where we observe the largest consistency effect. As contrast increases, response bias decreases as does the influence of the surround. This implies that subject response could be more dependent on surround information when the low-contrast feedforward input is particularly difficult to distinguish. As this information becomes more visible, the context has less of an influence.

It is tempting to view this contextual bias as a potential confound, however, it could perhaps indicate a mechanism which the visual system employs when the feedforward information is especially degraded. The reliance upon the information in the surround could be scalar, dependent upon availability of the feedforward input. If subjects were responding wholly according to the surround, performance accuracy in the Consistent condition would be near 100% and accuracy in the Inconsistent condition would be 0% (as the target does not match the surround in this condition, therefore response to the task would always be incorrect). Similarly, we would find evidence of extreme bias in both the hit rate and false alarm rate which would be evident in the sensitivity and response bias measures. We do not see clear evidence of a strong response bias in our results, and therefore it cannot fully explain our findings. Rather, it is likely the pattern of results we observe is due partially to a bias and partially to a consistency effect.

2.5.1.2 The consistency effect persists around threshold contrast-level

We observe the same pattern of results when we test both detection rate and recognition accuracy around threshold contrast level. We limited contrast levels to around the perceptual threshold (based on detection rates shown in the Consistency Effect Experiment and the Simplified Experiment) and found no significant effect of consistency on subjects’ ability to detect the low-contrast information. Detection rates for degraded information in the target region were similar for consistent and inconsistent contexts, implying that predictable low-contrast information is no more detectable than unpredictable threshold-level information.

At these low contrast levels, we do however find a significant advantage in being able to detect low-contrast information in the target region when there is no information in the surround. As such, subjects have a significantly higher detection rate when there is no high-contrast contextual information present, compared to when there is information available in the surround, regardless of its consistency. In other words, high contrast information in the surround, regardless of its relevance, appears to interfere with the detection of low-contrast target information (contrasting findings from (Sasaki et al., 2006) who find that a degree of noise enhances detection). Such an effect could be due to a phenomenon known as visual crowding in which target information is difficult to perceive in the presence of nearby flankers (Ronconi et al., 2016). Visual crowding has been shown to occur in natural scenes (Levi, 2011) and its neural underpinning has been identified as early as V1 (Millin et al., 2014). It is however, a process thought to only interfere with recognition, and not detection (Levi, 2011; Ronconi et al., 2016) but there is some evidence of crowding effects of detection, modulated by complexity of the visual scene (Põder, 2008). Crowding effects are weak whenever the target stands out from within the stimulus array and strong when it forms a coherent texture with the flanking stimuli (Saarela et al., 2009), which could account for our finding here. However, as visual crowding effects are unusual within detection tasks, our result here could simply reflect a more straightforward contrast effect. In the Target Only condition, the contrast level of the target region exceeds the contrast-level of the surround (as the surround is presented as a uniform grey image), making the target information more visible, by default. In the conditions in which there is information in the surround, the contrast-level of the target region does not exceed that of the surround and thus, the target scene is not as readily detected.

It is interesting to note that we only observe this effect in the Simplified Experiment, in which the contrast levels of the target images are presented around threshold contrast. We do not see any advantage of a lack of contextual information with a larger range of contrast levels, despite them remaining low in the Contrast Effect and Simplified Experiments. Perhaps the higher contrast levels used in these experiments are sufficient to detect a scene in the target region despite the cluttering effects of the high contrast surround, overriding the discrepancies between contrast levels of the two stimulus image elements. Although this finding may offer insight into such aspects of visual perception, it does not help our understanding of the distinction between predictable and unpredictable information processing. With no significant differences being observed between the consistent and inconsistent conditions, the data provide no meaningful effects of consistency on the detection of degraded feedforward information.

With regards to recognition around the threshold, we see the same pattern of response accuracy across consistency levels as observed in the experiments investigating a wider range of contrast levels, although in this experiment it doesn’t quite reach significance. Consistency between the surround and target regions enables greater performance in the recognition task than having either no information in the surround or inconsistent information in the surround. Thus, the effects of contextual modulation appear to withstand even when the feedforward input is very near threshold.

In terms of signal detection, we find a slight bias towards subjects reporting that the surround and target are consistent. This reduces as contrast increases and despite this, we still observe accurate sensitivity overall, which increases with contrast. We also observe increased sensitivity when the surround and target are consistent, and decreased sensitivity when they are inconsistent, when taking a general response bias into account. These results imply that subjects have a tendency to report that the target matches the surround, especially when the target region is most degraded. When contrast level increases, this tendency decreases, and overall sensitivity increases. This suggests that even around the perceptual threshold, where the contrast range is reduced, there could be multiple mechanisms at play in helping subjects process degraded feedforward information.

2.5.1.3 The consistency effect can extend beyond simple image features

We wanted to explore properties the consistency effect found in the recognition task. In The Consistency Effect Experiment, we find an advantage in task performance in all three of the consistency conditions. Subjects are better able to identify the content of the low-contrast target when surrounded by information that is consistent in terms of the task requirements. In The Consistency Effect Experiment, subjects were asked whether the target contained a natural or manmade scene. In each of the consistency conditions, when the target was natural, the surround was also natural and when the target was manmade, the surround was also manmade, however in two of the three consistency conditions, the surround and target images do not form parts of the same image. Despite this, we still observe a clear consistency advantage in these conditions.

In the Beyond Simple Image Features Experiment we wanted to see if this finding could be replicated around the threshold contrast-level. We find a significant effect of consistency, as in the Consistency Effect Experiment, whereby trials in which the surround and target were consistent either at the subordinate- or basic-category level resulted in higher recognition accuracy than trials in which there was no information or inconsistent information in the surround. As we replicate this finding despite the surround and target images never containing the same stimulus image, this suggests that the facilitatory consistency effect extends beyond the simple image feature of line continuation. When the context provided by the surround is useful for interpreting the low-contrast target information, without providing predictions about the actual stimulus image in the target, we see an increase in performance accuracy.

Using a signal detection approach, we find an increase in sensitivity within all consistent conditions, when a general response bias is taken into account. When consistency between the surround and target regions is used to measure response bias and d Prime, we still find subjects are able to accurately determine the content of the target region overall. Again, we find that response bias reduced as contrast increases and sensitivity also increases.

To experimentally address the potential for response bias a-priori, rather than account for this analytically, a control experiment was conducted within the lab. The experimental design of the Response Bias Experiment minimised the possibility of response bias influencing performance by including a 2-AFC task which meant subjects could not be biased toward selecting the category based solely on the context. Results of this experiment support the notion of consistency-related contextual facilitation on a recognition task; however, this is only found in instances in which the context contains the same image as the low-contrast target region, suggesting low-level image features may be a driving mechanism behind the consistency advantage, especially at low-contrast levels.

Taking both experiments into account, it is unclear whether the consistency effect is driven by low-level features as evidence for both is provided. The experiments differ in terms of methodologies and analyses, and it is therefore difficult to draw direct comparisons.

2.5.1.4 Altering spatial frequencies modulates this effect

Having found a fairly robust consistency effect across experiments, we then explored different properties of the feedforward and feedback information to see how limiting certain aspects (namely spatial frequency information) of the stimulus image would influence the ability to interpret the low-contrast target information.

Firstly, we limited spatial frequencies in the low-contrast target region, further degrading the feedforward input. Performance accuracy in this experiment was generally lower than previous experiments, indicating that further degradation of the feedforward signal increased task difficulty.

We find a significant difference between presenting the target region in either low or high spatial frequencies. Recognition accuracy was higher in the low spatial frequency condition, implying that removing high spatial frequency information in the target region had less of an impact on task performance than removing the low spatial frequency information. Current literature proposes that low spatial frequency information is extracted first from the visual scene and is used to prime context frames and associations in the higher visual areas which are fed to the lower visual areas and integrated with a more bottom-up, high spatial frequency driven scene extraction (Bar, 2003). The notion of low spatial frequency information driving an initial response capturing the gist of a scene which is then corroborated with finer details provided by higher spatial frequencies is well supported (Kauffmann et al., 2014; Schyns & Oliva, 1994). The finding of this experiment would therefore lend evidence to this notion and indicate that this spatial frequency distinction remains when information is ambiguous or difficult to perceive.

Furthermore, additional degradation of the feedforward information eliminates the consistency effect that we have reported thus far. Consistency between the surround and target is no longer influential on task performance when spatial frequency of the low-contrast information is limited. We also no longer find a significant increase in performance with increased contrast level. There are no significant interactions between the predictors; importantly, this suggests that the consistency effect also diminishes within each spatial frequency. In other words, we no longer find a beneficial effect of consistent contextual information (and vice versa for inconsistent contextual information) within either spatial frequency condition, as well as overall, collapsed across both spatial frequency conditions.

This pattern of results observed in the accuracy data are also reflected in the Signal Detection measures. We find only very slight evidence of a response bias towards reporting that the surround and target information are consistent, despite the notion to adhere to such a bias in the lower contrast levels in previous experiments which is thought to be a reflection of task difficulty. Yet here, further degradation of the target region does not result in such a pronounced bias, suggesting contextual information is not as influential when spatial frequency information is manipulated. We find a general response bias toward responding that the target region is a mountain, which is more pronounced in the higher spatial frequency condition. The response bias only reduces to around chance for the condition in which there is no information in the surround.

We find increased sensitivity when low spatial frequency information is preserved in both of the Signal Detection Analyses. This persists when both a general response bias and a bias toward responding to the surround are accounted for. As with the accuracy data, we do not find an influential effect of context in this experiment.

Results therefore indicate that reducing parameters of the feedforward input even further than previously tested not only increases task difficulty but also eliminates the influence of context. We find that removing high spatial frequency information is less of a detriment than removing low spatial frequency information, however, performance is driven mainly by the degraded feedforward input as top-down influence driven by the context no longer exerts a behavioural effect.

In a separate experiment, we limited the spatial frequencies in the full-contrast surround and the spatial frequency information in the low-contrast target region remained intact. We find a significant effect of spatial frequency, however, in contrast to the previous experiment, we find that removing low spatial frequencies has less of and influence on task performance than removing high spatial frequencies. Performance was therefore higher when the surround was limited to high spatial frequencies, versus low. This extends beyond accuracy as we find increased sensitivity when high spatial frequency information is maintained in the surround. This opposes the views of the general aforementioned literature in which low spatial frequencies appear to drive top-down predictions. Our results however imply that maintaining high spatial frequency information in the context is more central to performance than low spatial frequency information. Unlike when the spatial frequencies are limited in the target region, in this experiment we still observe a pattern within the consistency conditions, with consistent information enhancing performance. This consistency effect, however, does not quite reach significance.

We find a bias across all trials to report that the target region contains a mountain, but this is less pronounced when the surround is shown at high spatial frequencies. We also find evidence of a slight bias toward reporting that the surround and target are consistent in this experiment, which is more prominent in the high spatial frequency condition. Despite this, we find higher sensitivity in the high spatial frequency trials, independent of the response bias toward the surround. We also find higher sensitivity when the surround and target are consistent, but this does not reach significance. Taken together, the results imply that despite degradation of the contextual information, subjects still use this to interpret the low-contrast target region and that it still exerts an effect at the consistency level, to a certain extent. In terms of spatial frequency however, it appears that high spatial frequency information in the context is more pertinent in determining the content of low-contrast information.

2.5.2 Key findings and how they fit within the current picture

With this experimental series, we aimed to investigate how top-down, and bottom-up information interacts to process degraded visual input. In doing so, we manipulated different aspects of each information stream as well as investigating different types of processing. We mainly focused on manipulating the reliability of the top-down information by modifying the consistency of the context. This bottom-up input was degraded to varying degrees to further explore this effect. We also investigated other low-level properties of both the top-down and bottom-up signal. Exploring these facets of visual processing will give insight into the underlying processing streams and how these ultimately influence behavioural response.

Our results provide evidence for a facilitatory effect of consistent top-down predictions on the processing of degraded feedforward information. Not only that, but we also find that inconsistent predictions hamper the ability to interpret low-contrast information. This adds to the literature surrounding the effect of context on object identification, with support of a bi-directional influence of context in recognition performance (Bar, 2004; Bar et al., 2006; Fenske et al., 2006). While some of this effect may be explained by the influence of a response bias (whether implicit or explicit) toward the content of the context, such a bias cannot wholly account for the effects. Evidence of such a bias is observed mainly in the lower contrast levels and decreases as the target region becomes more visible which is accompanied by an increase in sensitivity.

The fact that subjects are still able to accurately determine the content of the target region when any bias is accounted for (evident through sensitivity), could suggest that rather than performance being confounded by a response bias, a bias toward the surround could in fact be a mechanism in itself which is used when the feedforward information is sufficiently degraded. When more feedforward information is available (as contrast increases), subjects may respond less according to the surround and instead use the contextual cues to distinguish the content of the degraded input. Perhaps top-down influences are less relied upon when feedforward information is sufficient (Bar et al., 2006). As such, a response bias would instead reflect a separate (or conjunctive) process that is utilised under certain circumstances. It is also important to note that what we refer to as a response bias (where subjects tend to report that the target is the same as the surround) could in fact instead reflect a perceptual bias. According to Raslear (Raslear, 1985), a perceptual bias results from perceived changes in the stimulus whereas a response bias reflects an induced response preference. That is, rather than the task difficulty at the lower contrast levels making subjects more likely to respond according to the surround due to perceptual uncertainty (i.e., choosing to respond according to the surround as the target region is indistinguishable), it is possible that the contextual information dominates the perceptual processing meaning subjects exhibit a perceptual bias consistent with the information in the surround.

In terms of bottom-up influences, we find a robust effect of contrast across our experiments. Increasing the contrast of the target region and therefore the availability of the bottom-up input, by definition, results in increased performance both in terms of accuracy and sensitivity. The influential effect of context on the processing of this degraded, bottom-up input is evident around the perceptual threshold; implying the that top-down signals resulting from contextual information can help process bottom-up input that is not always perceptually visible. Manipulating the low-level property of spatial frequency within the low-contrast target region highlights the role of low-spatial frequencies in discerning ambiguous feedforward information, independent of available contextual information.

The role of bottom-up and top-down processing was explored within two different aspects of visual processing: detection and recognition. While we find an influence of context on subjects' ability to recognise a degraded scene, we do not find a similar effect in detectability of said scene. That is, the content of the context does not appear to have a beneficial or detrimental effect on whether low contrast information can be identified. Instead, we find evidence to suggest that merely the presence of context may reduce subjects' ability to detect low-contrast information around the perceptual threshold. This can perhaps be explained by attributes of the visual system such as perceptual crowding (Levi, 2011). However, real-world scenes are naturally 'crowded' in terms of the vast array of elements, contours, textures etc. that they contain. It would seem problematic if a visual system built for processing complex scenes were to function more proficiently when contextual information is eliminated. That being said, we rarely have the need to process information degraded to such an extent as the target scenes presented in this experimental series. It is more likely that the effect we observe is simply due to a contrast effect in which the target region having higher contrast when surrounded by a blank grey image makes it easier to detect than when surrounded by high contrast information.

With such diverging findings between different task goals (detecting versus identifying a scene), it is important to look at the differences between the two aspects of visual processing within the visual system (Straube & Fahle, 2011). These distinctions highlight the different roles of detection and identification in visual processing and may account at least partially for the differences we observe.

Of course, our results may also partially reflect limitations within the experiment itself and not just limitations in the interpretation of the results. Firstly, sample sizes differed considerably between experiments (ranging from N=28 to N=7) and overall were low for gold-standard psychophysical experiments in order to obtain the desired power (Brysbaert, 2019). Including more subjects would be first priority in re-investigating the effects found within the experimental series as well as ensuring an equal number of subjects across each experiment.

Secondly, the dichotomous nature of the task (yes/no) response may limit conclusions that can be drawn. Including a confidence rating for example may allow for a more representative picture of subjects' detection/recognition ability to be drawn as it provides an indication of how certain subjects are of their response. Confidence ratings are often used within SDT-based analyses, however including such ratings increases the number of thresholds that are calculated which again increases with the number of categories. A larger data set (i.e. more trials) would need to be collected in order to capture each threshold reliably (Selker et al., 2019). Our experiment already contained a number of conditions so perhaps this is more feasible if we limited our conditions (in the interest of time for participants' motivation and attention) in future.

Thirdly, the difficulty of the task itself may be limiting the results. We may find more or less of an effect by increasing the visibility of the target region so that performance was more likely to remain above chance level. This however may negate or compromise the need for contextual information if the feedforward signal is too visible and mask any effects the reliance on feedback signals has. However, it could still be informative to test the effect at higher contrast levels (albeit before reaching the perceptual ceiling) to investigate whether there is a parametric shift in any effect when more feedforward information is available.

Finally, it could be argued that the placement of the occluded region may influence the effects found and applying this paradigm using different areas of the visual field as the target region may reveal different findings. There is a lower field bias for object perception and hand actions towards objects (Rossit et al., 2013; Schmidtmann et al., 2015) whereby subjects are better at perceiving objects presented in the lower visual field relative to the upper visual field. Scenes however encompass the whole visual field (Groen et al., 2017). Performance asymmetries have been found along the vertical and horizontal meridians (Abrams et al., 2012; Corbett & Carrasco, 2011) with optimal performance being just below fixation. As scenes subtend the periphery, it may be worth exploring occluding other regions of the visual field. This has been tested previously in our lab (Morgan et al., 2019) with the upper quadrant being occluded without much impact. With this in mind, we reasoned that key scene information is more likely to be in the lower visual field (as upper is most likely to contain sky) and therefore this is most appropriate to occlude within this study.

One way to account for our data (which we do not do here) could be by applying a computational model which would predict recognition based on the core features of bottom-up input (including parameters such as stimulus visibility and stimulus features available (for example type of spatial frequency information) and top-down signals (including strength of the top-down signal, features contained within that signal and its consistency with the bottom-up signal). Response bias could also be implemented as a feature of the model to account for any tendencies to respond in a certain way. The model would then predict subjects' ability to recognise degraded feedforward information depending on the parameter inputs estimated by the model. By carefully weighting the relative contribution of these features the model could account for the findings we report here, however difficulties lie within the contradictory findings of our data compared to the literature (for example our finding that the contributions of HSF information outperformed LSF input) therefore inconsistencies such as this would need to be considered when applying such a model.

Despite these limitations, our results give insight into how bottom-up and top-down signals influence the processing of degraded visual input. We find evidence of contextual modulation within the visual system by finding support for the role of feedback in the interpretation of feedforward input and highlight an interaction of the two processing streams when both types of information are available. This interaction ultimately manifests as a behavioural output with top-down influences altering subjects' behavioural response to feedforward information.

The findings of this experimental series, however, cannot give insight into the underlying neural processes involved in such interactions. There is a wealth of literature investigating the role of feedback at the neuronal level and exploring the contextual effect we find here within the visual cortex itself forms the basis of Chapter 3 of this thesis. Ultimately, we interact with the environment on a behavioural level, which reflects the output of underlying neural processes and interactions, so the findings here provide key insight into the way in which the visual system operates by highlighting the consequences of such interactions and prompts the exploration of this effect at the cortical level.

3 Chapter 3

3.1 Abstract

The influence of top-down predictions on the processing of visual input is well documented (Gilbert & Li, 2013) but the way in which they exert an effect at the cortical level is less clear. That is, the way in which top-down and bottom-up inputs are integrated at the neuronal level is still under debate. Some accounts postulate that the representation of predictable information is enhanced (Phillips, 2017) whereas others suggest that unpredictable information drives a greater cortical response (Rao & Ballard, 1999). Here, we aim to test whether top-down predicted information can amplify dim, feedforward input.

To achieve this, we employed a partial occlusion paradigm in which we presented a series of natural scenes with the bottom-right corner shown at low contrast, thus degrading the feedforward input to a region of early visual cortex. The content of the occluded region either matched or did not match the rest of the scene, manipulating the consistency (and therefore predictability) of the context. Using 3T fMRI we investigated the cortical response to such contextual modulation in early visual cortex, while simultaneously testing subjects' ability to recognise the content of the occluded region, psychophysically.

We tested this effect in areas V1 and V2 using both univariate and multivariate approaches. We also ran two experiments in which the contrast of the occluded region was either shown at 50% threshold-level or slightly higher at 75% threshold-contrast. In terms of BOLD response amplitude, we do not find a strong effect of consistency in V1 but do find evidence of the presence of feedback signals. V2 results suggest an enhancement of inconsistent information, suggesting higher areas may be more susceptible to visual scene inconsistencies. These findings are generally supported at the multivariate level and provide an insight into the cortical interaction between top-down and bottom-up signals.

3.2 Introduction

Imagine yourself looking through a pane of frosted glass. You would be able to make out the shape and form of large objects, but the quality of the visual information is drastically reduced, and many details of the scene are compromised. Nevertheless, you would still be able to infer what is on the other side to a degree, despite the level of visual information you are provided with being severely impaired compared to typical viewing conditions. How does the brain achieve this? How can we still interpret visual information when it is significantly degraded?

In this project we explore aspects of visual processing related to these questions. Specifically, we want to assess whether top-down predicted information can amplify dim, feedforward input. The dim feedforward input refers to the aforementioned frosted glass effect. If visual input is degraded in some aspect, we want to investigate how top-down predicted information influences the processing of such information.

Predictions in general serve an important purpose in the brain's function. Research has found evidence of feedback across all sensory modalities. In the visual system, feedback signals have been shown to play both a modulatory and driving role in mechanisms such as gain control and predictive processing (see (Muckli & Petro, 2013) for details). In auditory processing (Bonte et al., 2006) outline how feedback signals may be used to extract meaning from complex streams of sound. Using fMRI, Grabenhorst and colleagues (Grabenhorst et al., 2008) demonstrated both behavioural and neural modulation of taste and flavour perception using words to influence the cognitive labels associated with the stimuli and thus in turn, influencing the top-down signals. Manita and colleagues (Manita et al., 2015) found evidence for a neural circuit containing both top-down and bottom-up components in mouse somatosensory cortex. Finally, with electrophysiological recordings, Zagha (Zagha et al., 2013) demonstrate cortico-cortico feedback connections between the motor and somatosensory cortices, with motor cortex activity influencing network activity in the somatosensory cortex within their study. This widespread evidence of feedback signalling across the cortex suggests that feedback is a critical aspect of cortical function. The general consensus on the role of predictions is that they serve to facilitate the processing of the outside world by speeding up/reducing the processing of predictable stimuli while disambiguating information that is not so clearly distinguishable.

The role of feedback in visual processing is largely undervalued in classical, hierarchical models of visual processing (Grill-Spector & Malach, 2004b; Riesenhuber & Tomaso Poggio, 1999). They propose a sequential, feedforward flow of information and neglect to sufficiently acknowledge the top-down counter processing stream. However, architecturally, feedback connections outnumber feedforward connections within cortex (Clavagnier, Falchier, & Kennedy, 2004; Larkum, 2013) and cognitively, top-down processing has been shown to play a crucial role in both scene and object recognition. Historical studies have shown an advantage for possible (versus impossible) inter-object relationships within scenes (Hock et al., 1978) as well as a disadvantage for inconsistent object-scene relationships. This interplay between object and scene relationships is reinforced in later studies by Oliva & Torralba (Oliva & Torralba, 2006, 2007) who emphasise the importance of a scene-centred approach to investigating object processing (as objects do not appear in isolation in a real-world environment). They postulate that a scene-centred approach constrains local feature analysis, enhancing object recognition within natural scenes. In 2018, Lauer and colleagues (Lauer et al., 2018) conducted a series of experiments looking at the role of context in both scene and object recognition and conclude that a scene's low-level features contribute to the processing of scenes in real-world environments again highlighting an interplay between object and scene processing. The notion of a top-down influence on object and scene recognition led Bar (Bar, 2003) to propose a mechanism for how this mechanism may function for object recognition. Bar suggests that a coarse representation of an object is quickly extracted and triggers top-down expectations of the most likely interpretation (thus limiting the number of representations that need to be considered). These are then integrated with the bottom-up signal until an agreement is met regarding the object's identity. There is a wealth of literature on the role of top-down processing on object and scene recognition (see (Bar, 2004) for a review). An understanding of the mechanisms behind counter-current processing is crucial in understanding visual processing as a whole.

Predictions are formed from past experiences; knowledge gained from previous exposure to similar situations is used to guide perception and behaviour within the current visual environment. Predictions can also be formed from current contextual cues. Contextual information can be used to draw on prior experience and make predictions about the upcoming sensory information. Behaviourally, consistent contextual information has been shown to facilitate object recognition (Bar, 2004; Bar et al., 2006) and disambiguate ambiguous objects (Bar & Ullman, 1996). Contextual modulation therefore appears to play a key role in predictive processing and allows elements of the visual scene to generate predictions about the scene identity by combining top-down and bottom-up information.

In a broad sense, contextual modulation flexibly fits prior knowledge to current circumstances (Phillips, Clark, & Silverstein, 2015) and is a process thought to underlie several cognitive functions, meaning its utilisation is likely to be widespread throughout the cortex. How this process is achieved, however, is still not clear and there is evidence supporting several contrasting cortical theories which underlie contextual modulation.

A significant discrepancy between different theories of cortical function lies with how top-down and bottom-up information is integrated. Theories such as Adaptive Resonance Theory (ART), (Carpenter & Grossberg, 2016; Grossberg, 2013) propose that when incoming data matches expectations, there is an amplification and prolongation of neural signal. In other words, when sensory information matches our predictions, the signal is amplified and when the expectations and feedforward input do not match, the signal is suppressed. Contextual modulation also plays a central role in the theory of coherent infomax (Kay, Floreano, & Phillips, 1998; Kay & Phillips, 2011) which works on the basis that a key objective of cognitive processing is to maximise the amount of coherent information that is transmitted. This is achieved through amplification of coherent signal and suppression of incoherence, echoing the concept behind ART.

A proposed neural mechanism underpinning these two theories comes from apical amplification (Phillips, 2017). The concept behind apical amplification goes beyond viewing neurons as integrate-and-fire point processors, in which firing output depends on the trade-off between excitatory and inhibitory input and whether a threshold is exceeded as a result. Instead, apical amplification functions on the basis that neurons (specifically layer 5 Pyramidal neurons) have two integration sites which interact to determine the cell’s output. A somatic integration site receives driving input for feedforward signals, and an apical integration site for modulatory input (Phillips, 2017). As these two sites are within the same cell, the process serves as an intracellular mechanism underlying theories which implicate the amplification/suppression of relevant/irrelevant signals, respectively.

An intracellular process in which apical amplification can be achieved is through backpropogation-activated coupling (BAC), (Larkum, 2013). Larkum and colleagues outline the two integration sites in the soma and apical dendrites and provide intracellular evidence of the modulatory effect of the apical tuft. Feedback signals providing context and predictions target these apical dendrites via calcium spiking. When initiated, this causes a high frequency burst of action potentials in the soma. Therefore, feedback can drastically increase the cell’s overall output, further enhancing the notion that the sensory brain functions in a concurrent manner.

While the theories and mechanisms outlined above suggest that coherent contextual information is what amplifies the neuronal response to visual stimulation, there is a wealth of literature to suggest the opposite. One such framework is predictive coding. Predictive coding is built on the notion of Bayesian inferencing and stipulates that disparities between predictions and feedforward input are what is passed on to the next stage of the hierarchy (Rao & Ballard, 1999). Compatible feedforward and feedback signals are ‘explained away’ at the earliest opportunity so that only information which may require further processing to establish its meaning reaches the next stage of processing. These so-called prediction errors are addressed at every level through integration of top-down and bottom-up signals, with the ultimate goal of free energy minimisation (Friston, 2005). Empirical evidence for predictive coding accounts of processing come from many lines of research (e.g., Alink et al., 2010). All ultimately provide evidence for a reduction in neural activity when stimulation is consistent with predictions and an increase in signal when there are discrepancies, contrasting what is proposed by theories explained by apical amplification. Furthermore, there is also evidence to suggest that rather than an enhancement or decrease in neural signal modulated by expectations, top-down information may serve to sharpen the neural representation of predictable information (Abdelhack & Kamitani, 2018; Kok et al., 2012).

While these theories contrast in their propositions as to how the neural signal should respond to consistent versus inconsistent contextual information, they do share a common feature in that feedback signals play an important role in visual processing and have the ability to modulate the feedforward signal. This modulation is achieved through the amplification and/or suppression of select signals depending on their coherence with the bottom-up input. The exact nature of this interaction between the two processing streams is yet to be fully established however this study aims to help to shed some light onto the neural consequences of bottom-up top-down information integration.

Within this study, we want to address the key question of whether top-down predicted information can amplify dim, feedforward input. To do this, we employ a partial occlusion paradigm coined by Smith & Muckli (Smith & Muckli, 2010). Occlusion is a regular occurrence during natural scene viewing. Scene elements and objects within a scene are rarely viewed in isolation of each other, instead they overlap and occlude portions of the scene from view at any given time point. Occlusion during experimentation of visual processing allows feedback signals to be isolated in a region of visual cortex by eliminating direct feedforward input to the corresponding area of the visual field.

In Smith & Muckli's paradigm, they occlude one quadrant of the visual scene and are able to decode information about the surrounding scene in occluded visual cortex. This implies that when feedforward information is missing, feedback signals carry scene information to the subsequent region of visual cortex. We aim to adapt this paradigm so that our occluded region contains degraded feedforward input. Our stimuli will also consist of natural scenes however, the occluded region will also contain low-contrast scene information that either matches or does not match the main image, thus manipulating the predictability of the top-down signals.

We will investigate the cortical implications of such an occlusion using 3T fMRI to gauge whether BOLD response to predictable information in the visual cortex is amplified relative to non-predictive signals. fMRI is a useful tool in mapping neural correlates of brain function and is sometimes considered advantageous over measures of behaviour itself (Logothetis, 2008). Applying such a technique is therefore a natural progression of this paradigm. Furthermore, we will apply a pattern classification technique (MVPA) so that we can also assess the response pattern to predictable versus non-predictable visual information. Differences that exist within BOLD response to different stimuli may be missed if the focus is limited to univariate responses such as response amplification. Multivariate analysis such as MVPA may reveal differences in response patterns which are not possible to discern at the univariate level.

Finding evidence of either an amplification or a decrease in response to predictable or unpredictable information or a difference in the pattern of response will facilitate our understanding of predictive visual processing at the neuronal level and help map the connection between behavioural and cortical responses.

3.3 Methods

3.3.1 Paradigm

We used the same partial occlusion paradigm used in the psychophysical experiments outlined in Chapter 2 (Smith & Muckli, 2010) in which part of the visual scene is degraded in order to investigate how top-down predictions influence the processing of degraded feedforward information. For paradigm details, see section 2.3.2 and Figure 2.3 (a) for a depiction of the basic stimulus features.

The behavioural studies previously outlined inform the motivation for this Chapter. They provide tentative evidence of a behavioural effect of context and therefore we are keen to explore if there are any neural underpinnings of a contextual effect within this paradigm. As this is an fMRI experiment, we looked at both the behavioural and cortical consequences of these consistent or inconsistent feedback signals. We used performance accuracy and sensitivity measures for the psychophysical responses and both univariate and multivariate measures of BOLD signal change to capture the neural response pattern.

3.3.2 Stimuli

To generate our stimuli, we used two images depicting natural scenes: one of a beach and one of a mountain. These were chosen from the same image database as the stimulus images used in the psychophysical experiment series (Walther et al., 2009) and passed through a Spatial Envelope Model (Oliva & Torralba, 2001) to control for image features.

We created the stimulus images using MATLAB (R2016a). The image for the surround was greyscaled and cropped to comply with the screen resolution inside the scanner (1024 x 768 pixels). A second image was then selected for the target region. The contrast level of the target image was adjusted to the threshold contrast-level of the subject. In the Lower-Contrast Experiment we used each subject’s 50% threshold contrast level (the contrast level at which the subject could detect a target region scene 50% of the time) and in the Higher-Contrast

Experiment, we used each subject’s 75% threshold contrast level. The two images were then combined so that the low-contrast target region occupied the bottom right corner (24.8° x 18.7° VA) of the stimulus image and the remaining three quadrants contained the full-contrast surround. A fixation cross (0.1° x 0.1° VA) was overlaid in the centre of the image to direct subjects’ gaze and this stimulus image as a whole was presented during the task trial, subtending 47.5° x 36.5° VA.

The content of both the surround and target regions was manipulated across four experimental conditions: Consistent, wherein the surround and target regions contained the same stimulus image; Inconsistent in which the two regions contained different images; Target Only, in which there was no image displayed in the surround, only a low-contrast image in the target region; and finally Context Only, which contained no image in the target region but a full-contrast image in the surround. We also included a Null condition, which contained no image in either the surround or target region.

3.3.3 Subjects

We recruited subjects via a participant subject pool advertised through the University of Glasgow, Institute of Neuroscience and Psychology intranet. All subjects had healthy or corrected vision. During the recruitment process, subjects were sent an MRI pre-screening checklist to ensure they were eligible to take part in an MRI experiment. Once confirmed, we arranged the scanning session.

Upon arrival at the scanning session and before entering the scanner, subjects gave informed consent to take part in the experiment and completed an MRI safety checklist. To comply with MRI safety guidelines, subjects were required to remove any metal such as piercings or hair ties before entering the scanner. Participants also wore MRI-compatible clothing provided by the University during their scan. Subjects were paid for their participation (regardless of whether they completed the whole session) at a standard rate of GBP 10.00 per hour.

A total of 74 participants were recruited (37 for each experiment). Data from 35% of the sample were excluded at different stages of acquisition or analysis for reasons such as incomplete datasets or excessive head movement within the scanner. As such, results are reported from forty-eight subjects, twenty-four subjects from the Lower-Contrast Experiment (16 female, M age = 22.13, SD = 4.12) and twenty-four from the Higher-Contrast Experiment (13 female, M age = 22.92, SD = 3.94).

3.3.4 Data Acquisition

Data were acquired on a Siemens Tim Trio System 3T MRI scanner, located at the University of Glasgow, Centre of Cognitive Neuroimaging (CCNi). We used a 32-channel head coil to acquire anatomical and functional data of the brain. Anatomical data were acquired using a T1-weighted high-resolution Gradient-Echo (GE) sequence, imaging the whole brain with 192 volumes and an isotropic voxel size of 1mm. Functional images were acquired using an Echo-Planar-Imaging (EPI) sequence to measure the BOLD response in line with the functional task. We acquired 18 interleaved axial slices, positioned optimally over the occipital pole, to ensure sufficient capture of primary visual cortex. Slice thickness was 3mm with an inter-slice gap of 0.3mm and a voxel size of 3mm isotropic. We used a TR of 1000ms and a TE of 30ms. Matrix size was 70 x 70, and we used a flip angle of 62°. Field of view (FOV) was 210 x 210mm. Each experimental functional run consisted of 757 volumes and the retinotopic mapping run used the same scan parameters but contained 795 volumes. Experimental stimuli were presented using Presentation software (Neurobehavioral Systems, Inc., Berkeley, CA; www.neurobs.com) and projected to a projector screen located within the scanner.

3.3.5 Procedure

Upon arrival to the scanning session, we carefully explained the task requirements and obtained written consent and completion of the MRI safety checklist. Subjects had chance to ask any questions about the task itself or the process of an MRI experiment. We provided MRI compliant clothing. If subjects had vision corrected by glasses, we also provided MRI compatible goggles in which lenses matching individual subject prescriptions can be inserted. We were careful to ensure subjects have no metal on their person before entering the scanner.

When entering the scan room, subjects were given earplugs to protect from the noise of the scanner and asked to lie on the scanner bed in a head-first supine position. Cushions or inflatable pads were placed around the head to minimise head movement during the scan. Subjects were given a response box in their right hand and an emergency buzzer in their left hand. The head coil was placed over their head followed by a mirror to enable them to view the projector screen. Once subjects were comfortable, they were moved inside the scanner and the lights were turned off. From the control room, we were able to verbally communicate with the subjects through a microphone and speaker system. We requested subjects to test the emergency buzzer before the scanning session began.

Prior to the main experiment, subjects completed a behavioural pre-screening test whilst in the scanner (Figure 3.1), to ascertain their threshold contrast level. This is performed during the anatomical scan to minimise time spent inside the scanner. The pre-screening test employed the same paradigm as the main experiment, except with only three experimental conditions: Consistent, Inconsistent and Target Only. The two image exemplars used to create the stimuli (one of a beach and one of a mountain) were different to the two used in the main experiment, to minimise any practice effects.

In the pre-screening test, the target region in each condition was shown at seven contrast levels, ranging from 0.6%-1.8%, in increments of 2%, in the Lower-Contrast Experiment, and from 1.0%-2.2% in the Higher-Contrast Experiment, also in increments of 2%. Stimulus images were presented for 1000ms flickering at 4Hz. Following each trial, subjects performed a detection task in which they were asked whether the target region contained a scene. Responses were two-alternative *(yes/no)* and were indicated by button press on the response box. See Figure 3.1 for details of the pre-screen experiment.

Subjects remained in the scanner while the pre-screen data were analysed. This identified each subject’s detection threshold (50% in the Lower-Contrast Experiment or 75% in the Higher-Contrast Experiment) which was used to determine the contrast level for the target region in the main experimental stimuli. In this time, we reminded subjects about the task requirements, either verbally or by writing instructions on the screen.

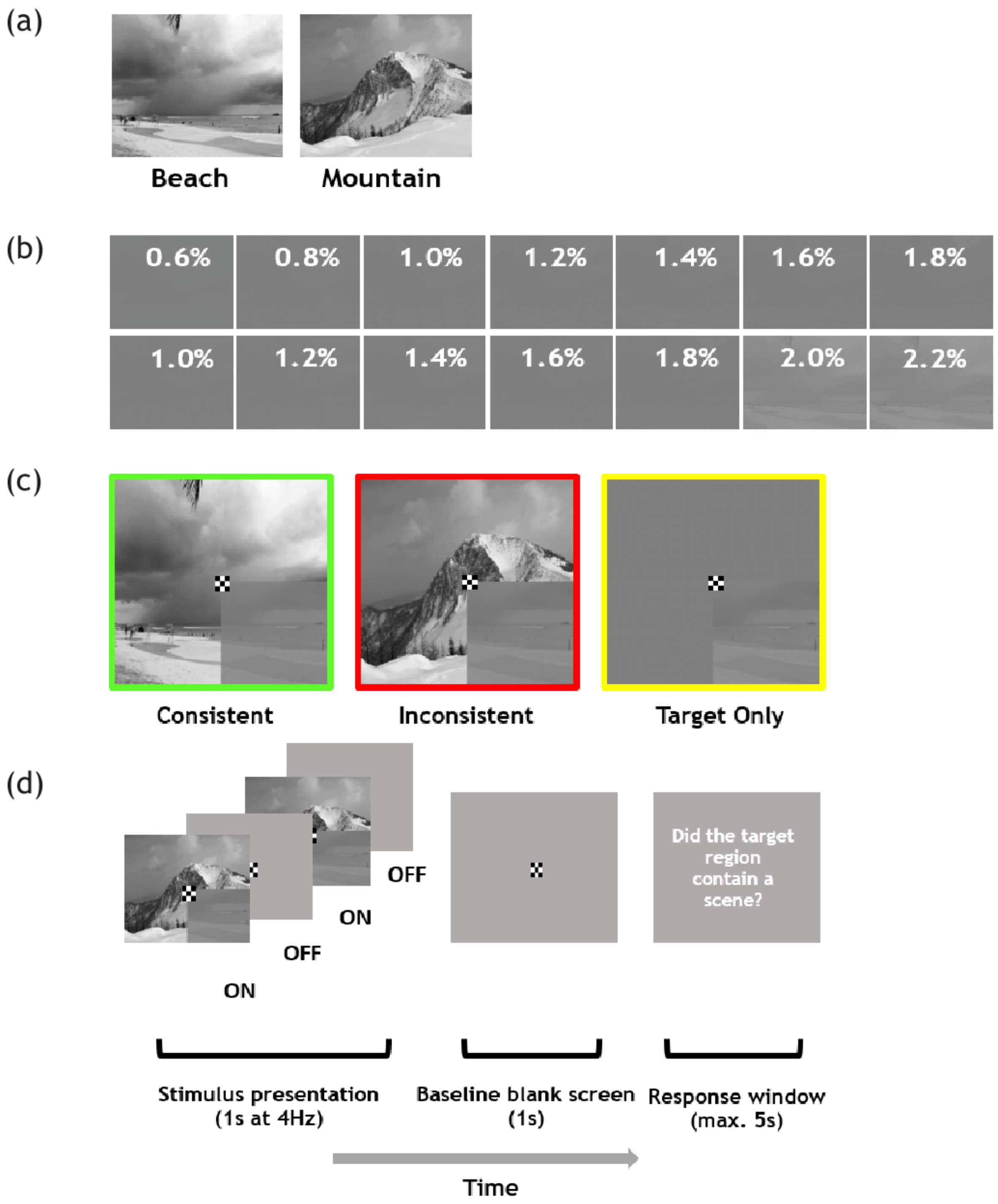


Figure 3.1 Pre-screen experiment details. (a) Images used to create the stimuli; one beach and one mountain. We used different image exemplars in the pre-screen experiment to the main fMRI experiment to avoid any practice effects. (b) Contrast levels used for the target region in the Lower-Contrast Experiment (top row) and the Higher-Contrast Experiment (bottom row). These were used to determine subjects' 50% and 75% detection thresholds, respectively. (c) Experimental conditions: Consistent – surround and target region contain the same image; Inconsistent – surround and target region contain different images; Target Only – no information is shown in the surround. Note that the target region is shown at high contrast here for demonstration purposes. (d) Trial sequence. The stimulus image is flashed on/off at 4Hz for 1s followed by a baseline blank screen before the task question (*did the target region contain a scene?)* is displayed until subjects respond either yes or no (or for a maximum of 5s if there is no response).

The main fMRI experiment (Figure 3.2) consisted of four experimental runs, each containing 47 trials and lasting approximately twelve minutes. Each 14s trial consisted of 3s stimulus presentation (stimulus on) followed by 3s of a blank screen (stimulus off) then a 2s response window in which the task question was presented and finally a 6s inter-stimulus-interval (ISI) before the next trial began.

Each experimental condition (bar the Null condition) had two sub-conditions pertaining to each of the two stimulus images. For example, the Consistent condition contained both Consistent Beach stimulus images and Consistent Mountain stimulus images. Thus, each sub-condition is repeated five times in each run, plus five Null trials in which there is no image in either the surround or the target region. The first two trials were repeated at the end of the run, to account for the haemodynamic saturation experienced at the beginning of the run. The first two trials were therefore removed in the analysis stage.

At the end of each run, we performed mapping of the target, border and surround regions, in turn. Contrast-reversing flickering (4Hz) checkerboard stimuli were shown, occupying the target (21.8° x 15.4° VA), surround (47.5° x 36.5° VA) and border regions, mapping the border between the surround and target with a width of 3.2° x 3.2° VA. Following the four experimental runs, we performed retinotopic mapping, which consisted of one run, also lasting approximately twelve minutes. A combined eccentricity and polar angle (PA) mapping stimulus (again contrast-reversing checkerboards, flickering at 4Hz) traversed the visual field, mapping the PA and eccentricity retinotopic organisation of each subject’s visual cortex.

Mapping the cortex using tailored mapping stimuli alongside retinotopic mapping enables us to determine the cortical location of the target region within the V1, V2 and V3 boundaries of each subject.

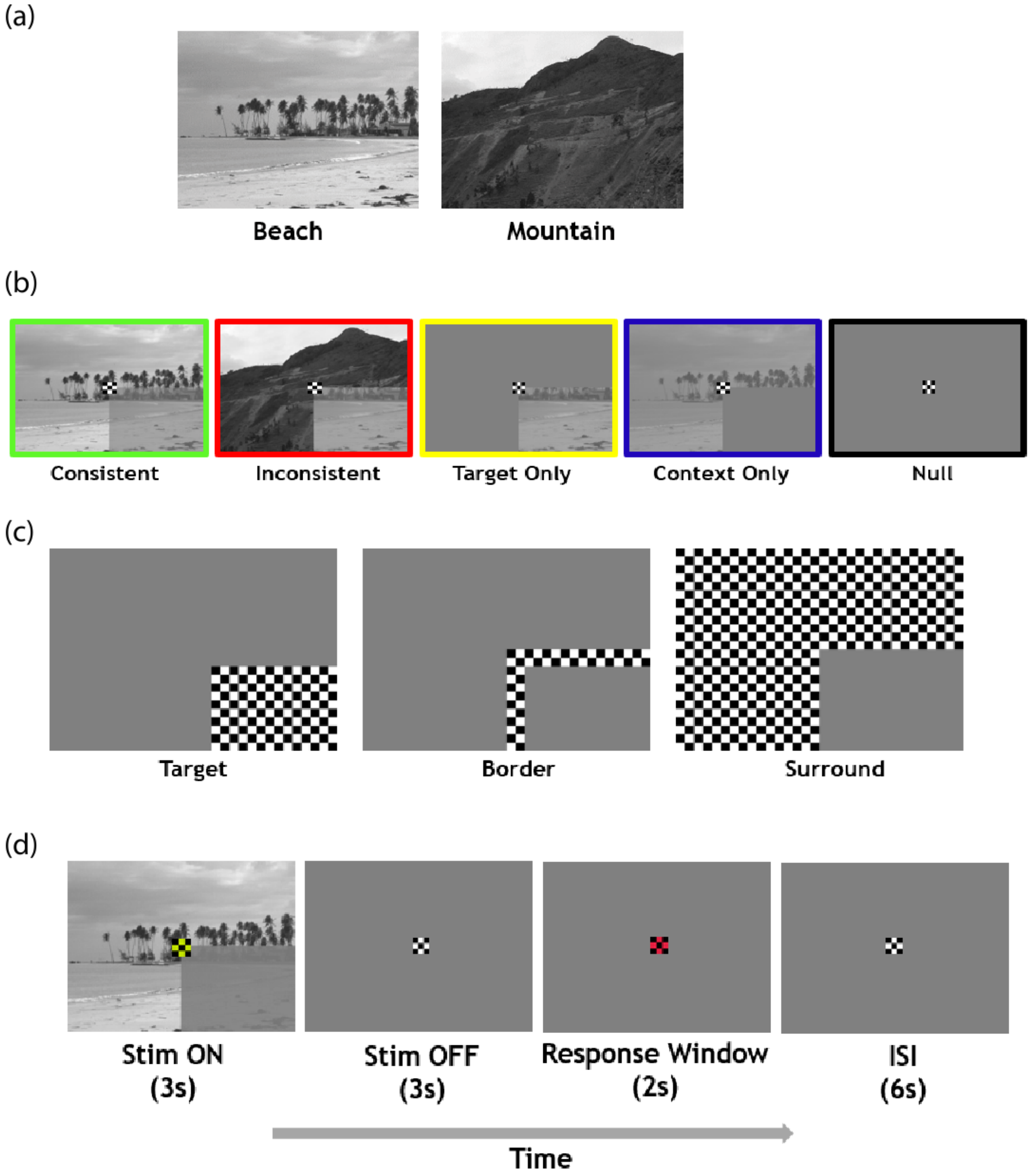


Figure 3.2 fMRI experiment details. (a) The image exemplars used in the main fMRI experiment. One beach and one mountain which were different to those used in the pre-screen experiment. (b) The experimental conditions: Consistent – the surround and target regions contain the same image; Inconsistent – the surround and target regions contain different images; Target Only – Information is displayed in the target region only and the surround is left blank; Context Only – information is shown only in the surround; Null – no scene information is shown in either the surround or target regions. (c) Mapping stimuli. Flickering checkerboards used to map the target region (left), border (middle) and surround (right). Responses to these stimuli were used to identify the ROIs for each subject in early visual cortex. (d) Trial sequence. The stimulus image was displayed for 3s (with a green fixation cross) followed by a 3s baseline blank screen (with a white fixation cross). The fixation cross then changed to red for 2s in which subjects were required to respond with beach or mountain (pertaining to the target region scene). Finally, a 6 second inter-stimulus-interval (ISI), with a white fixation cross, allowed for some haemodynamic response recovery before the next trial commenced.

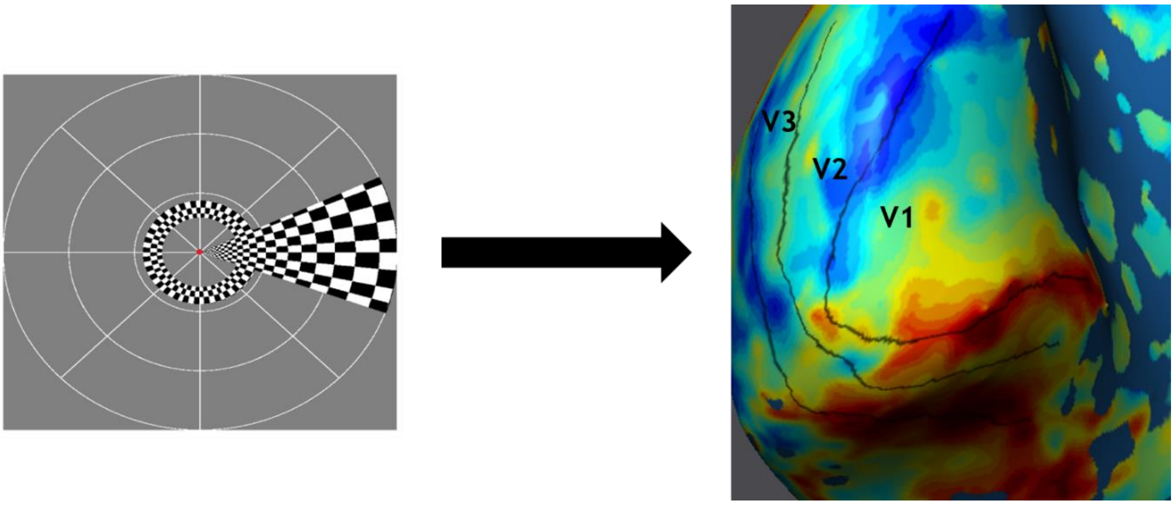
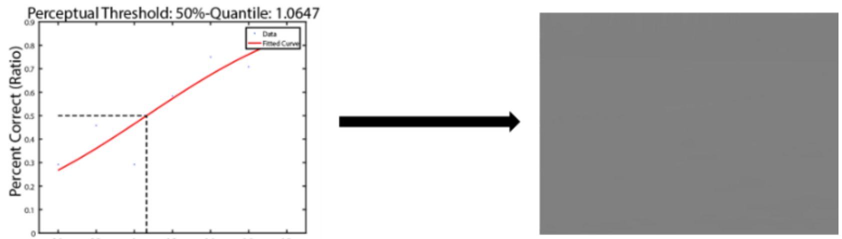


Figure 3.3 Retinotopic mapping exemplar. Example retinotopic mapping stimulation (left), in which a combined polar-angle and eccentricity mapping stimulus rotates and expands periodically to map the visual field using high-contrast flickering checkerboards. An example polar angle retinotopic map (right), demonstrating areas V1, V2 and V3 in the visual cortex, here located in a left hemisphere occipital pole/cortex. This map is obtained by a flickering checkerboard wedge stimulus which rotates around the centre of the visual field while subjects maintain fixation.

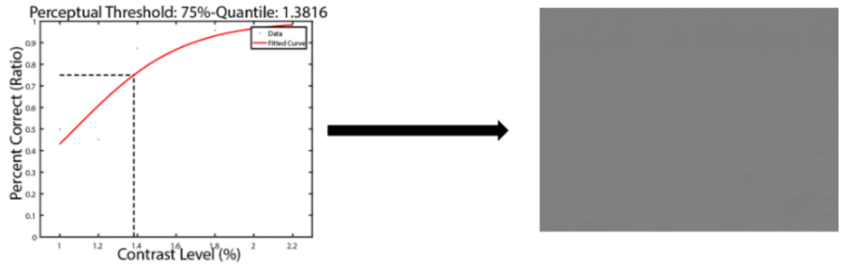
3.3.6 Analysis

3.3.6.1 Pre-screen

Analysis of the pre-screen task response was performed while subjects remained in the scanner using MATLAB (R2016a). We fitted a Sigmoid function to each subject’s responses, averaged across all consistency conditions, plotting their detection rate as a function of contrast, to obtain their threshold contrast level (the value of which, dependent on experiment). This contrast level was then used to generate the stimulus images required for the main experiment. See



Sigmoid fit of perceptual threshold (50%)



Sigmoid fit of perceptual threshold (75%)

Figure 3.4 Determining stimulus contrast level. We fitted a sigmoid function to subjects' detection-rate data from the pre-screen experiment. This enabled us to determine the point at which subjects could detect a stimulus either 50% of the time (top row – used for the Lower-Contrast Experiment) or 75% of the time (bottom row – used for the Higher-Contrast Experiment). The left-hand images show example Sigmoid functions, and the right-hand images show stimuli at the resulting contrast level.

3.3.6.2 Psychophysical Task

Responses to the psychophysical task are recorded and used to calculate percentage accuracy in each consistency condition. That is, subjects' ability to correctly identify the low-contrast information in the target region, within each condition. Performance accuracy (as a percentage) for each consistency condition was plotted as a function of contrast level. Although accuracy can be used to quantify task performance, it does not take into account the underlying components of a behavioural response. We therefore also analysed the behavioural data using the principles of Signal Detection Theory (Swets et al., 1978), which provides an independent measure of sensitivity while simultaneously accounting for any response bias.

We used the principles of SDT to split the responses into hits, misses, false alarms and correct rejections using the subject response and content of the target region (see Table 3.1 for details). These values were then used to calculate the hit and false alarm ratio (number of hits/number of trials in which the target and surround are consistent or number of false alarms/number of trials in which the surround and target are inconsistent). These ratios were then in turn, used to calculate d-prime (d'), a sensitivity measure. Sensitivity determines how effectively subjects can distinguish one category of stimulus over another while accounting for potential bias, so is therefore more informative than using accuracy as a measure of performance.

Surround	Target	Response	Category
Beach	Beach	Beach	*Hit*
Mountain	Mountain	Mountain	
Beach	Beach	Mountain	*Miss*
Mountain	Mountain	Beach	
Beach	Mountain	Beach	*False Alarm*
Mountain	Beach	Mountain	
Beach	Mountain	Mountain	*Correct Rejection*
Mountain	Beach	Beach	

Table 3.1 SDT response categorisation. How hits, misses, false alarms and correct rejections are determined for the psychophysical data depending on the consistency between the target and surround regions as well as the subject response.

3.3.6.3 Anatomical and fMRI Data

Anatomical data were analysed using BrainVoyager (version 21.2 and Qx) (Brain Innovation, Maastricht, Netherlands). Each subject's 3D scan images were passed through a brain extraction tool and corrected for inhomogeneities in the image intensities to maximise segmentation of white and grey matter. Anatomical data were then aligned along the AC-PC plane and co-registered with the functional data using boundary-based registration (BBR). Boundary-based registration uses the WM/GM borders to align the functional and anatomical data by maximising the boundaries between the tissue types (Greve & Fischl, 2009).

The anatomical data were also used in conjunction with the functional data to create high-resolution volume time courses in the visual cortex as well as define a volume of interest (in which to perform the statistical analysis). An inflated 3D mesh cortical representation was then created in which the functional and retinotopic mapping data were overlaid to allow the regions of activation and boundaries of early visual areas to be visualised.

Analysis of the fMRI data is also performed using BrainVoyager (versions 21.2 and Qx) and began with pre-processing. Data from each run were 3D motion corrected (using trilinear/sinc interpolation), slice scan time correction was applied using cubic spline interpolation as well as temporal filtering with a GLM-Fourier high-pass filter with a cycle of 6 sines/cosines. Alignment between functional runs was correlated using Pearson's r correlation and adjustments are made to ensure all functional runs are well aligned (>0.9). Anatomical data were co-registered with the functional data (using BBR) and a 3D mesh was created in which statistical maps could be overlaid.

We defined individual subject Regions of Interest (ROIs) for both V1 and V2 (where possible) by fitting a GLM to a conjunction contrast of the mapping trials. We are interested in testing for a contextual effect in early visual cortex and therefore any activation beyond V1 and V2 was not analysed. V3 was not included due to the limited number of voxels activated (due to the relative cortical size of V3) which would make multivariate analysis difficult. We contrasted the target versus border and target versus surround mapping blocks and used the activity which persisted in both of these contrasts to define our ROIs. This is a fairly conservative approach to ROI definition but was employed to minimise the risk of spill-over activity due to extra-classical RF properties that exist within visual cortex (Angelucci & Bressloff, 2006). This resulted in a total of 1041 voxels (mean voxels per subject= 43) in V1 and 818 voxels (mean voxels per subject also = 43) in V2. These numbers are fairly low due to the conservative nature of our contrast.

For the statistical analysis within the ROIs, we applied a deconvolution analysis due to its efficacy in analysing rapid event-related designs (Hinrichs et al., 2000). Due to the closely spaced events in such a design, the haemodynamic responses to each trial overlap considerably. These must be deconvolved in order to ascertain the response (defined as the activity in relation to the 3s of stimulus 'on' time) to each predictor. Predictors were defined in the experimental protocol files and correspond the consistency conditions (trials) to a series of twenty (1s) time points across the haemodynamic response, commencing at stimulus onset.

We fit a GLM to estimate the haemodynamic response function for each predictor (corrected for multiple comparisons). These haemodynamic responses were deconvolved to give a BOLD response amplitude value in the form of beta-weighted estimates for each condition, at each time point. Jittered timing was not used within this design to allow for averaging across trials and on the reasoning that the deconvolution analysis still enables the HRF signal to be deconvolved. Six contrasts were defined in order to statistically compare response amplitudes between conditions (Consistent versus Inconsistent; Consistent versus Target Only; Consistent versus Context Only; Inconsistent versus Target Only; Inconsistent versus Context Only; Target Only versus Context Only). The Null condition was initially set to be used as a baseline however we wanted to measure the difference in signal in the Consistent versus Inconsistent directly rather than each relative to baseline, therefore it was not included in the analysis. Using visual inspection, we separated the HDRF into three time windows which we labelled the *peak* (8-11s post-stimulus onset), *post-peak* (12-15s post-stimulus onset) and *recovery* (16-19s post-stimulus onset) windows. Within each of these time frames, we ran the six defined contrasts within the subjects' ROIs, allowing us to thoroughly compare the effect of contextual consistency on the univariate BOLD response amplitude in early visual cortex. Group-level results are reported below. We firstly report BOLD response amplitude results for all trials to search for a neuronal effect irrespective of behavioural performance. Including all trials also allowed for a more direct comparison of univariate and multivariate analyses as the multivariate analysis required a sufficient number of trials which is not guaranteed if limiting the analysis to correct trials only (see below for details on the multivariate analysis). We then confined the analysis to correct trials only (i.e., we removed trials in which subjects responded incorrectly or did not respond). When a subject responded incorrectly, we do not know the reason why. Therefore, removing these trials removes instances in which the subjects were not fully concentrating on the task (which may dilute any effect). In other words, limiting the analysis to correct trials only may refine any influence of context on neuronal response.

We also examined the effects of context on BOLD response using multivariate analysis in the form of multivariate pattern classification (MVPA), (Haxby, 2012; Haxby et al., 2014). The classifier was coded in MATLAB as part of the original partial occlusion paradigm employed by Smith & Muckli (Smith & Muckli, 2010) and operates using a Support Vector Machine (SVM) algorithm which is trained to distinguish voxel response patterns. As our stimuli were comprised of two types of natural scenes (beach and mountain), the classifier aimed to discriminate between the response patterns in the target region pertaining to each image type.

The volume time courses for voxels within the defined VOIs are input into the classification analysis along with a design matrix which defines the predictor variables (outlining conditions and timings within each run). A GLM is then run to obtain betas and t-values, giving an estimate of the effect. Post-processing steps include excluding values with low signal change (as these are likely to represent noise). A dissimilarity matrix is then calculated to represent similarity between evoked responses to the predictors. Voxels at the border between the target and surround regions are then excluded to avoid any spill-over effects from surround region stimulation. The analysis then assigns and stacks betas to their corresponding conditions and labels these appropriately for classification.

We employed both cross-validation and cross-classification multivariate analysis techniques. In cross-validation, the classifier is trained and tested on the same condition in a leave-one-run-out manner. For each sub-condition (consistent beach, consistent mountain, etc.), the classifier was trained on identifying voxel response patterns to each image type in n-1 runs and tested on the remaining run. This was repeated until all runs were used as the testing run in turn.

In cross-classification, the classifier was trained and tested on different conditions. We firstly trained the classifier to identify the response pattern of the target region using the feedback signals only (the Context Only condition) and tested on all other conditions. We then repeated this process using the Target Only condition (i.e., purely feedforward information) followed by the Consistent condition for training. In both types of classification analysis, we average across both sub-conditions to give a classifier accuracy percentage for each consistency condition and ran the classifier for each individual subject before averaging to get a group-level indication of classifier performance. To test the classifier performance statistically, we used the Wilcoxon Signed Rank test to test for differences between conditions as well as to test whether classifier performance differed significantly from chance. This replicates a classification-based significance test performed by Morgan and colleagues (Morgan et al., 2019).

3.4 Results – Lower Contrast Experiment

3.4.1 Psychophysics

Subjects performed a behavioural recognition task while in the scanner. For each trial, they were required to identify whether the target region contained a beach or a mountain. Results reveal increased performance accuracy in the recognition task when the information in the surround was consistent with the target region. In other words, when both the surround and target regions contain a beach or both contain a mountain, subjects were more easily able to identify the content of the low-contrast target region (Mean=67.32%, SD=16.29%), than if there was no information in the surround (Mean=62.50%, SD=13.85%). Performance is lowest in trials in which the surround and target contain different images (Mean=52.51%, SD=18.45%). A repeated measures ANOVA reveals that these differences are significant, with a significant main effect of consistency being found using Greenhouse-Geisser corrections (F(1.527, 35.128)=6.935, p=.006). Post-hoc pairwise comparisons using Bonferroni corrections reveal significant differences between all consistency levels except the Consistent and Target Only conditions.

These findings imply that contextual information in the surround influences the ability to recognise low-contrast information in the target region, with consistent context enhancing performance and inconsistent context resulting in chance-level performance. However, analysis of the data using principles from Signal Detection Theory determines the underlying mechanisms of response bias and sensitivity which drive the behavioural output. Using SDT, we can investigate whether the behavioural effects we observe can be explained, at least in part, by a response bias toward the surround. That is, is the performance difference that we find between the Consistent/Inconsistent conditions due to subjects' tendency to base their response on the information shown in the high-contrast surround?

Accuracy as a measure performance is insufficient in determining the difference between a highly sensitive observer (i.e., someone who is accurately able to determine the content of the target region) and a highly biased observer (e.g., someone who has a strong tendency to respond according to the surround) – these two types of responders would both have high performance accuracy. SDT, however, is able to provide a measure of response bias (how likely subject are to consider the surround and target regions to be consistent) and sensitivity (a measure of performance, independent of a response bias. Hits, misses, false alarms and correct rejections were determined using the consistency between the surround and target regions alongside the subject response (see Table 3.1 for details).

Looking at response bias, a negative criterion value indicates that subjects are more likely to respond that the surround and target regions are consistent. Across subjects, we observe a criterion value of -0.2246, meaning subjects are slightly inclined to respond according to the surround. Despite this slight bias, we observe a sensitivity (dPrime) value above chance (0.5942, with 0 being chance level). These results imply that although subjects show a tendency towards the surround, when this is taken into account, they are still able to determine the content of the target region. The way in which the hit and false alarm rates are calculated using surround and target consistency does not allow us to determine these measures within conditions, and instead these values represent bias and sensitivity across all trials. Therefore, we are unable to infer whether sensitivity is indeed enhanced when the target and surround are consistent, but we are able to assume that the results do not wholly reflect a response bias. Figure 3.5 and Table 3.2 summarise the results of the psychophysical task.

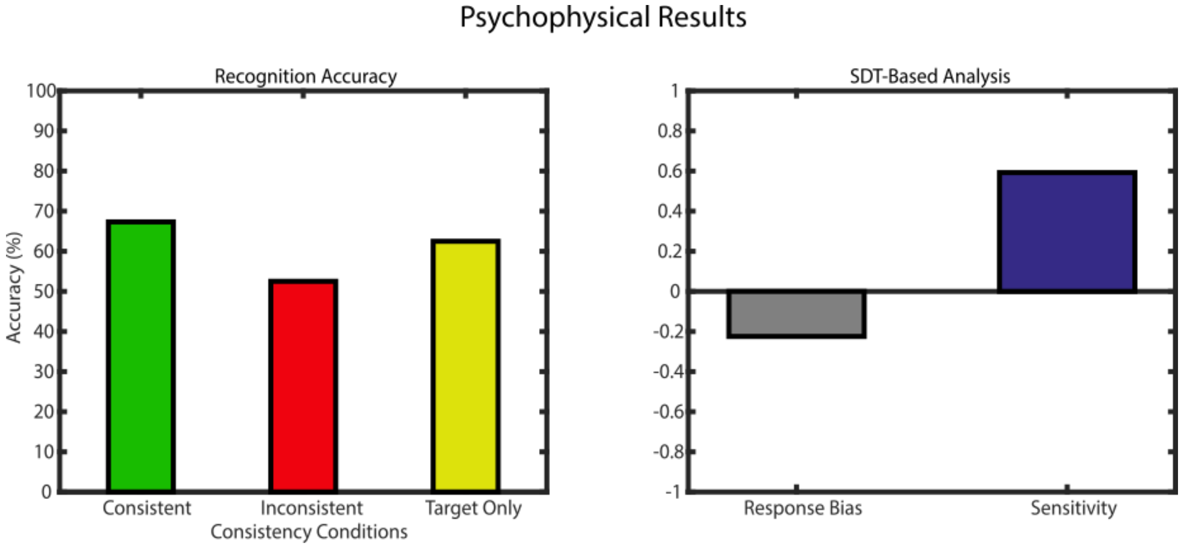


Figure 3.5 Psychophysical Results. The left figure depicts performance accuracy as a percentage within each consistency condition. The right figure displays the results of the SDT-based analysis in which a negative response bias indicates a tendency to respond that the surround and target region are consistent. Sensitivity (independent of response bias) is displayed in the right-hand column.

Hit Ratio	False Alarm Ratio	Response Bias (c)	Sensitivity (dPrime)
0.6732	0.4749	-0.2246	0.5942

Table 3.2 SDT results. Table displaying the SDT results with hit and false alarm ratio which are used to calculate response bias (criterion) and sensitivity (dPrime) values.

3.4.2 Univariate Analysis (V1)

3.4.2.1 All Trials

Turning our attention away from the psychophysical data, we then consider the *neural* response to low-contrast information when surrounded by consistent or inconsistent contextual information. We ran a Deconvolution Analysis on the region of interest corresponding to the stimulus target region in V1 which splits the BOLD response into twenty time points, post-stimulus onset. Upon visual inspection of the BOLD response at the individual and group level, we decided the peak BOLD response amplitude was between 8-11 seconds post stimulus onset. We therefore focused our analysis within this time window but to avoid neglecting the rest of the haemodynamic response, we also considered a post-peak time window (12-15 seconds post-stimulus-onset) and a time window during the haemodynamic recovery (between 16-20 seconds post-stimulus-onset). These time frames will be referred to as *peak*, *post-peak* and *recovery*, going forward. Within each of these, we investigate an effect in the BOLD response amplitude by reporting the beta estimate of response amplitude (which provides information about the amplitude of the effects itself) followed by an associated t-statistic (a directional measure offering substantiation of the effect estimate). It is important to consider both of these aspects of the effect (Chen et al., 2017).

In the peak response time-window, we find no significant difference in BOLD response amplitude between the Consistent and Inconsistent conditions (ß=.301 (SE=.282), t=1.067, p=.286), meaning we find that the area of cortex corresponding to the low-contrast target region has a similar peak BOLD response whether the contextual information in the surround is consistent or inconsistent. The Context Only condition (in which there is information in the surround but no scene presented in the target region) also elicits a similar response amplitude as both the Consistent and Inconsistent conditions and there is no significant difference between the Context Only and Consistent condition (ß=.282 (SE=.282), t=.999, p=.317) or the Context Only and Inconsistent condition (ß =-.019 (SE=.282), t=-.067, p=.926) within the peak response window. At the peak, the Target Only condition (the condition in which there is no information in the surround) elicits a highly significantly lower BOLD response than the other three conditions in which there is contextual information in the surround (ps<.001), implying feedback information from the surround drives the BOLD response in the conditions that include contextual information, at least in the peak response window.

In the post-peak time window, the BOLD response of the Consistent, Inconsistent and Context Only conditions decreases and the non-significant difference between the Consistent and Inconsistent conditions is maintained (ß =-.181 (SE=.283), t=-.640, p=.524). The Context Only condition elicits a significantly higher response than the Consistent condition within this time window (ß =-.64 (SE=.282), t=-2.265, p=.023) but not the Inconsistent or Target Only conditions (ps>.05). In contrast to the peak response window, the Target Only condition has a significantly higher BOLD response than both the Consistent (ß -.79 (SE=.283), t=-2.794, p=0.005) and Inconsistent (ß =-.610 (SE=.283), t=-2.157, p=.031) conditions. The response to the Target Only condition is also higher than the Context only condition in this time frame, but this does not reach significance (p=.596).

The response pattern observed in the post-peak time window becomes more apparent in the recovery window. The Consistent and Inconsistent conditions elicit the lowest response and there is no significant difference between the BOLD recovery amplitudes in these conditions (ß =.250 (SE=.284), t=.879, p=.379). The BOLD response in this time window in the Context Only condition is significantly higher than the Inconsistent (ß =-.778 (SE=.284), t=-.241, p=.006) but not the Consistent condition (ß =-.529 (SE=.284), t=-1.865, p=.062). The Target Only condition has the highest BOLD response in this time window, which is significantly higher than both the Consistent and Inconsistent conditions (ps=.001). The Target Only response is not significantly higher here than the Context Only condition (ß =.534 (SE=.284), t=1.876, p=.061).

In sum, it appears that during the peak, the BOLD response is driven by information in the surround as conditions which contain contextual information elicit a higher response amplitude than the condition in which no information is present in the surround. Conversely, following the peak, the low-contrast information in the target region drives a higher response but to a lesser extent than the distinction found at the peak. This is only the case in the purely feedforward driven (Target Only) condition as when this information is combined with feedback signals from the surround, as in the Consistent and Inconsistent conditions, the BOLD recovery amplitude is lower. Figure 3.6 depicts the BOLD response amplitude for each condition, as a function of time.

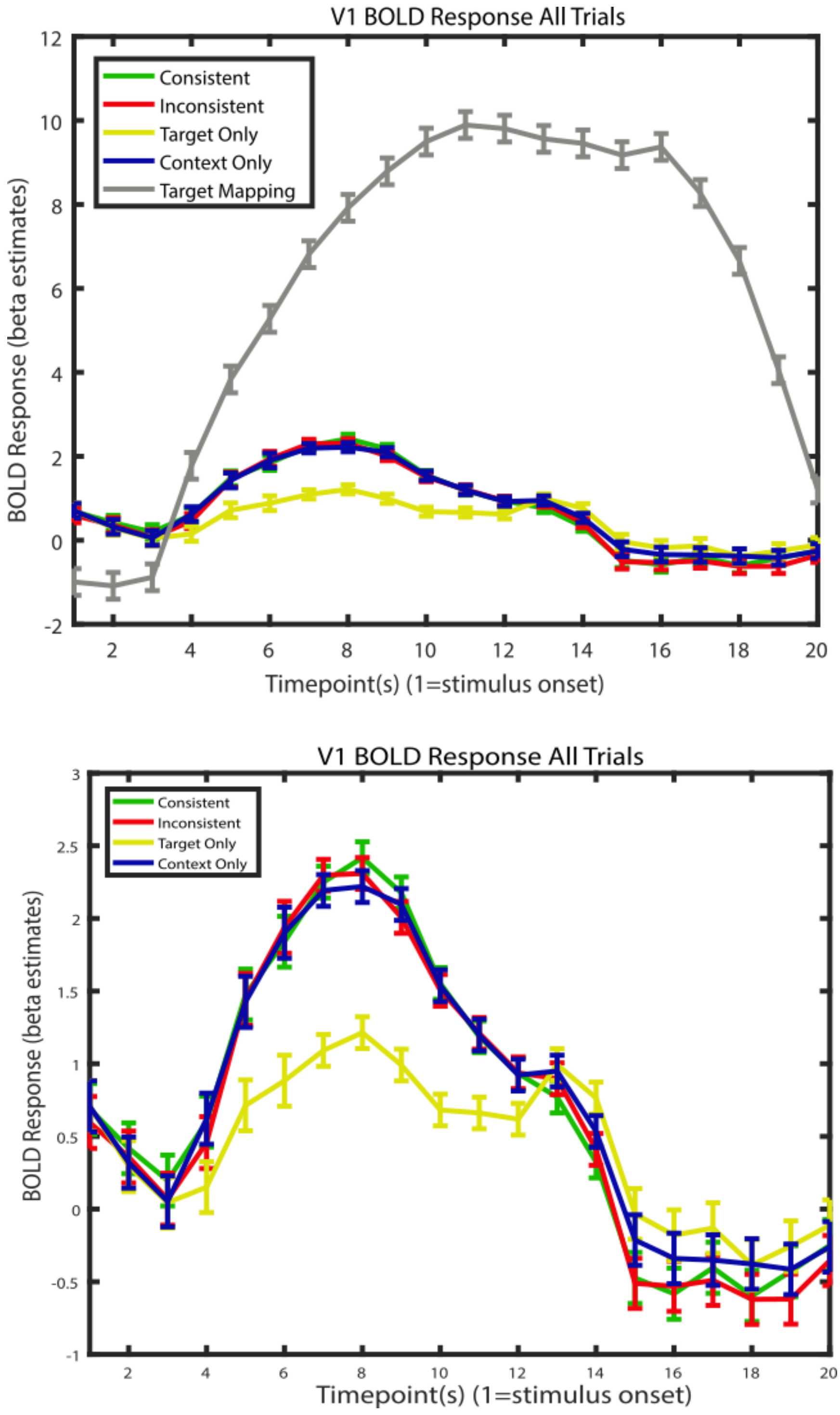


Figure 3.6 V1 BOLD response when all trials are included. Haemodynamic response as a function of time in V1 target region ROI for each experimental condition across participants (N=24) obtained from a deconvolution analysis. The top figure includes the target mapping condition to demonstrate the V1 response to high contrast stimulation. The bottom figure removes the target mapping condition to reveal differences between the experimental conditions. In both figures, the X axis corresponds to time in seconds post-stimulus-onset. The Y axis corresponds to the beta estimates of the BOLD response. Error bars represent Standard Error.

The magnitude of the BOLD response is important to note. The top figure of Figure 3.6 includes the cortical response to the target mapping condition. In this condition, the target region was stimulated by a high-contrast flickering checkerboard pattern and the surround was left as a blank grey screen. This condition can therefore be likened to a scenario in which the target region contained full-contrast information, without the influence of contextual feedback from the surround. We observe a considerably higher BOLD response to the target mapping stimulus than any of the four conditions of interest. Thus, illustrating that the contextual effects we are investigating are of smaller magnitude than if the target region were to be stimulated by non-degraded feedforward information.

It is also worth noting that in our first eight subjects, the fixation cross remained white for the duration of the trial (both stimulus presentation and blank grey screen during the ISI). We reasoned that as the target image was presented at threshold contrast level, there will be trials in which the subject does not consciously perceive any information in the target region. This could pose a problem during the Target-Only trials as subjects may not recognise a target region stimulus presence and thus, the processing of such trials could differ based on an attention-related effect. We therefore decided to proceed by including a fixation cross colour change when the stimulus was presented in all conditions, with the aim of reducing any impact of awareness (or lack thereof) in the Target Only condition. Figure 3.7 depicts the different BOLD response in the Target Only condition in the subjects without the fixation cross colour change (unaware - Target Only – U) and the subjects with the fixation cross colour change (aware - Target Only – A). When subjects are aware of a trial being presented, there is a considerably higher BOLD response than when subjects are unaware of a trial presentation, implying an attentional effect which drives a greater haemodynamic response.

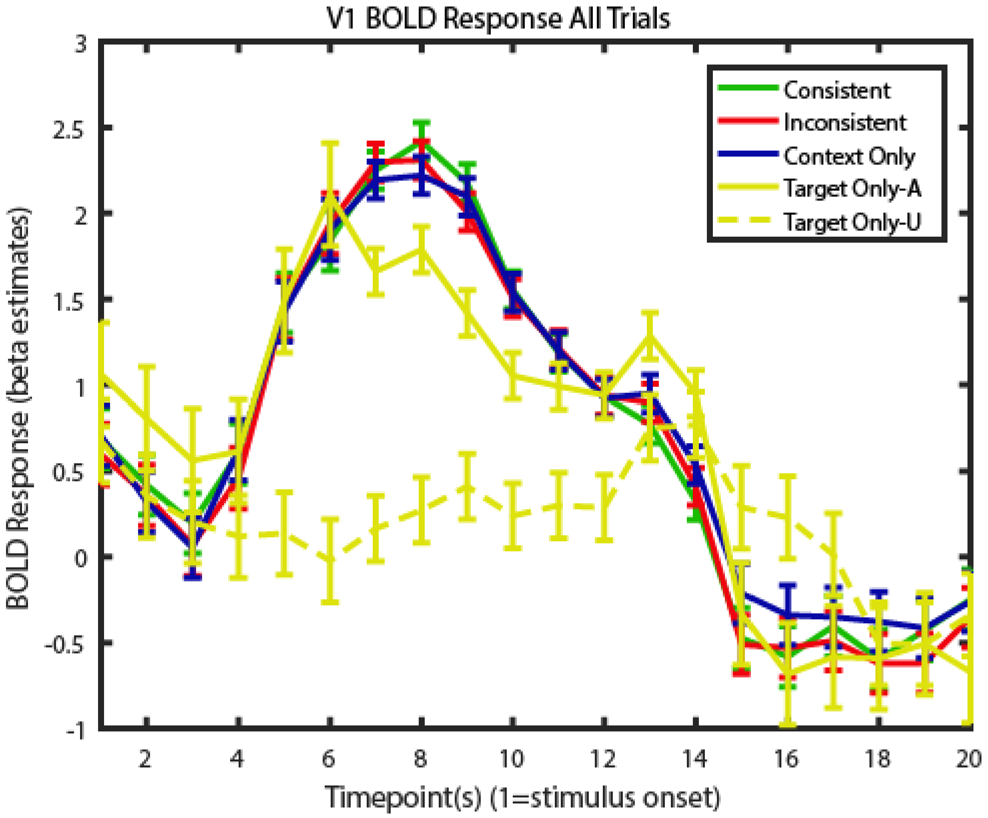


Figure 3.7 Comparing attentional effects. The haemodynamic response in V1 target region ROI across subjects, as a function of time. Focusing on the yellow lines which depict the Target Only condition in which there is no high-contrast information in the surround. The unaware subjects (Target Only – U) were presented with trials in which the fixation cross remained the same throughout the entire trial, meaning in the Target Only condition they were at least sometimes unaware that a stimulus had been presented. In the aware subjects (Target Only – A), we introduced a fixation cross colour change when a stimulus was being displayed, meaning they were then made aware that there was a stimulus present even in the Target Only condition. Error bars represent Standard Error.

3.4.2.2 Correct Trials Only

We then limit our interest only to trials in which the subjects responded correctly. At the peak time-window, we find no significant difference between the BOLD response amplitude in the Consistent and Inconsistent conditions (ß=.368 (SE=.371), t=.991, p=.321). Again, the Context Only condition elicits a similar response amplitude to the conditions in which there is scene information in both the surround and target regions (Consistent and Inconsistent conditions), however this is still significantly lower than the Consistent (ß =.702 (SE=.322), t=2.180, p=.029) but not the Inconsistent condition (ß =.334 (SE=.348), t=.961, p=.337), implying feedforward information does contribute to some of the peak BOLD response, especially when it is consistent with the information in the surround. The Target Only condition results in a significantly lower BOLD response amplitude than all other conditions which contain surround information, (ps<.001).

In the post-peak time window, the BOLD response for the conditions in which there is information in the surround (i.e., those which drive the highest peak BOLD response) decreases whereas the Target Only condition response remains fairly stable. As such, all conditions elicit a similar BOLD response amplitude in this time-window, and we find no significant differences between any conditions (all ps>0.05).

In the recovery time-window, we see a similar response pattern in the correct trials only as we do across all trials, with the conditions in which there is information in both the surround and target regions having more of an undershoot than trials in which information is absent in either the surround or target regions. Here, the response amplitude is significantly higher in the Target Only condition than both the Consistent (ß=-.929 (SE=.359), t=2.588, p.009) and Inconsistent (ß =-1.499 (SE=.381), t=-3.936, p=.001) conditions but is not significantly higher than the Context Only condition (ß =.493 (SE.335), t=1.473, p=.141). The response amplitude in the Context Only condition is significantly higher than the Inconsistent (ß1.006 (SE=.35) t=-2.873, p=.004) but not the Consistent (ß=-.436 (SE=.325), t=-1.343, p=.179) condition. Figure 3.8 displays the BOLD response in the correct trials only.

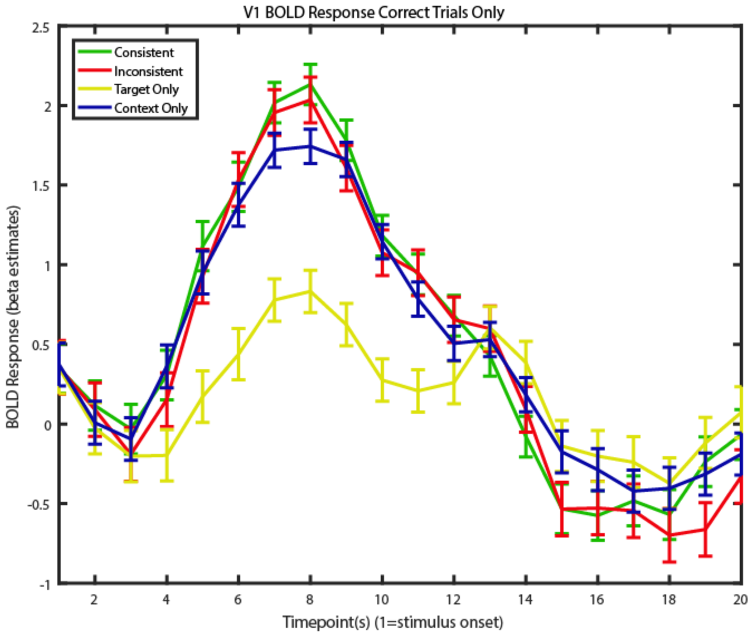


Figure 3.8 V1 BOLD response for correct trials only. Haemodynamic response in the target region of V1 for each consistency condition, for trials in which the subjects responded correctly. The x-axis displays time in seconds post-stimulus-onset and the y-axis displays the beta estimates of the BOLD response. Error bars represent Standard Error.

To sum, much like the response profile across all trials, in trials in which the subject responded correctly, the peak BOLD response appears to be driven by contextual information in the surround. When surround information is missing in the Target Only condition, the BOLD response amplitude is significantly lower, however this could also reflect the lower contrast of the available information as well as the fact that less of the visual field is being stimulated in the Target Only condition. Following the peak of the BOLD response, the conditions in which more information is available to the observer (in the form of both feedforward information in the surround and feedback information in the target region) have a greater undershoot than conditions in which information is absent to some degree. The type of information present in the context does not appear to influence the BOLD response amplitude.

3.4.3 Multivariate Analysis (V1)

3.4.3.1 Cross-Validation

We also ran multivariate pattern analysis on the V1 ROIs. While different information in the surround may not influence the signal amplitude, the pattern of response may differ depending on the content of the context. In this cross-validation, we trained and tested an SVM classifier to be able to decode the information in the target region (either a beach or a mountain) within each of the four consistency conditions. The classifier was trained on n-1 runs within each condition and tested on the remaining run. This was repeated until each run was tested in turn and the results were averaged across runs. (See Methods section 3.3.6 for further details).

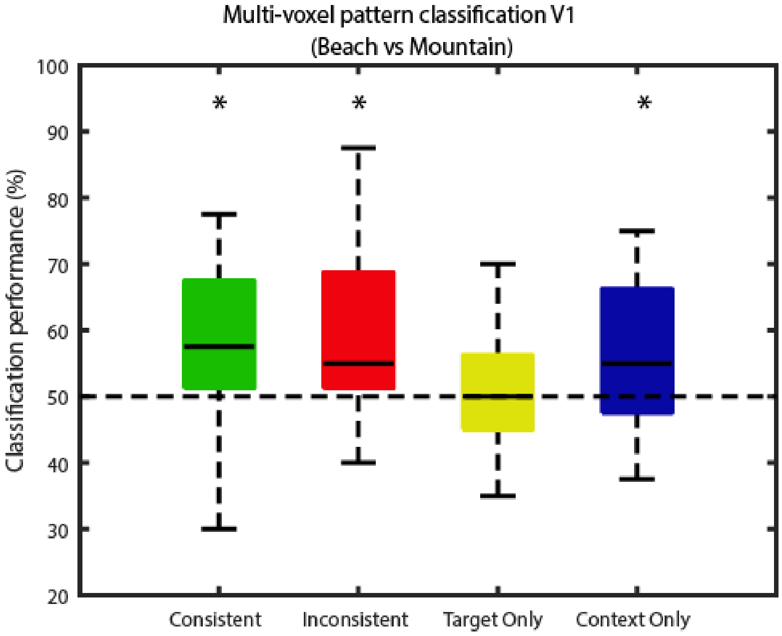


Figure 3.9 MVPA V1 cross-validation results. Results of a cross-validation analysis in which a classifier was trained on n-1 runs within each consistency condition to determine the content of the target region (beach versus mountain). Percentage classifier performance (y-axis) is displayed for each consistency condition (x-axis). The top and bottom edges of the box indicate the 75th and 25th percentiles, respectively, with the median denoted with a solid black line. The whiskers depict the range, excluding outliers (which are denoted by +). The horizontal dashed line indicates chance-level performance at 50%. Performance significantly different from chance is denoted by *.

Figure 3.9 displays the results of the cross-validation analysis. A Wilcoxon Signed rank test reveals the classifier is able to accurately decode the target region significantly above chance in the Consistent (Mean=57.92%, Median=57.5%, SD=11.34%), p=.004, Inconsistent (Mean=59.27%, Median=55%, SD=12.86%), p=.004, and Context Only (Mean=56.77%, Median=55%, SD=11.22%), p=.013, conditions. The classifier is unable to decode the target region in the Target Only condition (Mean=50.31%, Median=50.0%, SD=8.67%), (p=.861). Differences between the conditions in which the classifier can accurately decode are not significant (ps>.05), however the Target Only condition classifier performance is significantly lower than all other conditions (ps<.05). The classifier is therefore only able to decode when the stimulus condition contains information in the surround, implying that information from the surround is present in the BOLD response signal in the target region. The classifier loses its ability to decode the target region when no information is present in the surround and only low-contrast information is present in the target region.

3.4.3.2 Cross-Classification

Following cross-validation we then sought to test whether the classifier could cross-classify between conditions in the target region. We ran an SVM cross-classification analysis with three scenarios to test how accurately the classifier could distinguish between beach and mountain images between conditions. In the feedback cross-classification analysis, we trained the classifier on the Context Only condition which is akin to training purely based on feedback signals from the surround (as no feedforward information is present in the target region in this condition). Once trained on feedback information, we tested the classifier on being able to decode in each of the other conditions. Conversely, we then trained the classifier on the purely feedforward (Target Only) condition and tested on the other conditions. Lastly, we trained the classifier on the Consistent condition, which contains compatible feedforward and feedback signals and tested on each of the other conditions. Figure 3.10 summarises the cross-classification results.

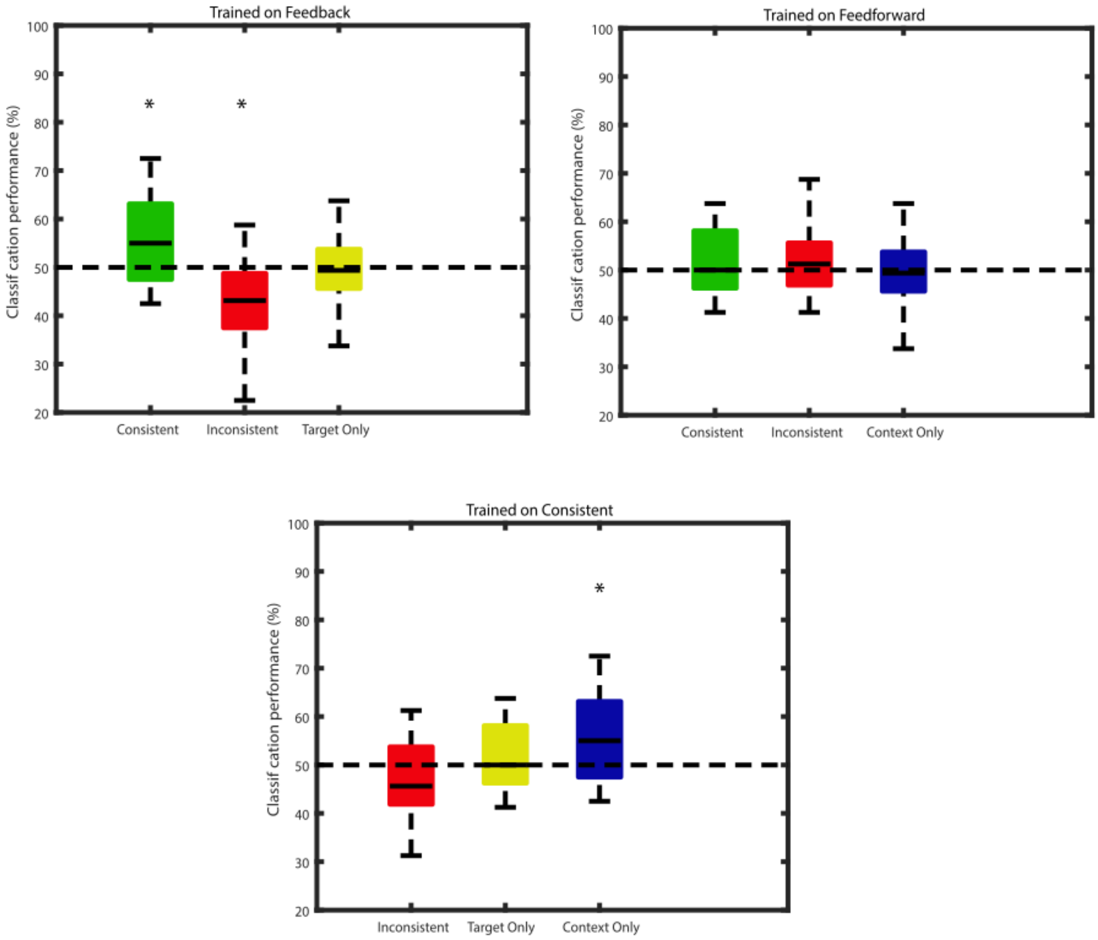


Figure 3.10 V1 cross-classification results. Results from the cross-classification MVPA analysis in which the classifier was trained on purely feedback (top), purely feedforward (middle) and consistent feedback and feedforward (bottom) information and tested on all other conditions in turn. Classifier performance as a percentage is plotted for each condition, reflecting the classifier's ability to determine the content of the target region as either a beach or mountain scene. The top and bottom edges of the box indicate the 75th and 25th percentiles, respectively, with the median denoted with a solid black line. The whiskers depict the range, excluding outliers (which are denoted by +). The dashed black line represents chance-level performance at 50%. Performance significantly different from chance is denoted by *.

When the classifier was trained on purely feedback information (Context Only condition), classifier performance was highest for the Consistent condition (Mean=55.89%, Median=55%, SD=8.96%), a Wilcoxon Signed rank test reveals this is significantly above chance (p=.009). Accuracy for the Inconsistent condition drops to significantly below chance (p=.001), with a mean accuracy of 43.18%, median of 43.13% and SD of 8.18%. When tested on purely feedforward information (Target Only condition), mean accuracy was 49.53%, with a median of 49.36% and SD of 6.34%. This is not significantly different from chance level (p=.772). Differences between all conditions were significant (ps<.05).

When trained on feedforward information (Target Only condition), the decoding accuracy does not differ significantly from chance in the Consistent (p=.250), Inconsistent (p=.279) or Context Only (p=.772) conditions. The Inconsistent condition results in a slightly higher decoding accuracy (Mean=55.93%, Median=51.25%, SD=6.78%) compared to the Consistent (Mean=51.78%, Median=50%, SD=6.52%) and Context Only (Mean=49.58%, Median=51.25%, SD=6.34%). The decoding accuracy did not differ significantly between conditions (ps>.05).

When trained on consistent feedforward and feedback information (Consistent condition), the classifier is only able to accurately decode above chance in the Context Only condition (p=.0092) with a mean decoding accuracy of 55.89% (Median=55%, SD=8.96%). Accuracy drops to around chance for the Target Only condition (Mean=51.77%, Median=50%, SD=6.52%) and slightly lower for the Inconsistent condition (Mean=47.08%, Median=45.63%, SD=7.85%), although neither of these differ significantly from chance level classification (p=.250 and p=.072, respectively). The classifier performed significantly worse in the Inconsistent condition than both the Target Only (p=.033) and Context Only (p=.011) conditions but the Target Only and Context Only conditions did not differ significantly from each other (p=.072) in classification performance.

To sum, when the classifier was trained on purely feedback information from the surround, it was only able to accurately decode the target region in the Consistent condition, when both feedback and feedforward information were compatible. Performance dropped to significantly below chance when the information in the surround and target regions were inconsistent and classifier performance was around chance level when tested on purely feedback information in the target region. When trained on purely feedforward information (low-contrast information in the target region) the classifier was not able to cross-classify significantly above chance when tested on any of the three remaining conditions. When trained on the condition in which both feedback and feedforward information is compatible, the classifier is only able to accurately decode when tested on information in the surround only.

3.4.4 Univariate Analysis V2

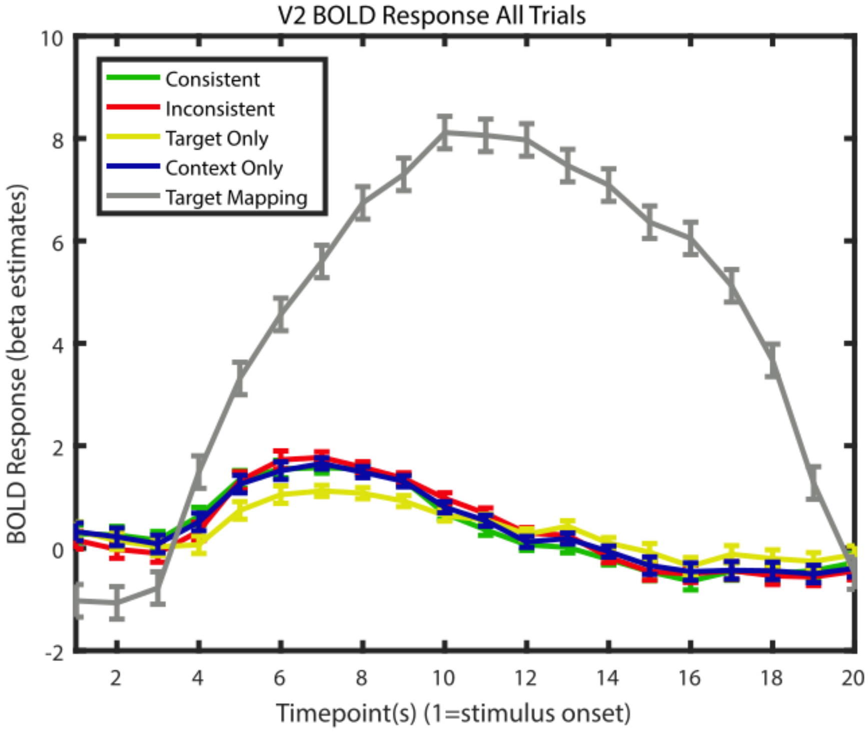
3.4.4.1 All Trials

We were able to adequately define V2 target region ROIs in twenty-two out of the twenty-four participants. Looking at BOLD response amplitude overall, the response is slightly lower in V2 than V1, for both the conditions in which the target region is shown at low-contrast and in the target mapping condition in which the target region contains high-contrast mapping stimuli.

Unlike in V1, we find a significant peak difference in BOLD response amplitude between the Consistent and Inconsistent conditions, with Inconsistent stimuli eliciting a significantly higher BOLD response than Consistent stimuli (ß=-.684 (SE=.281), t=-2.435, p=.0149) This distinction continues through the post-peak and recovery time frames but does not reach significance in either of these windows (ps>.05).

As in V1, the Target Only condition has a significantly lower peak BOLD response than the Consistent (=.724 (SE=.281), t=2.577, p=.009), Inconsistent (ß=1.409 (SE=.281), t=5.011, p<.001) and Context Only (ß=-.943 (SE=.281), t=3.354, p=.001) conditions. In the post-peak and recovery time windows, the Target Only condition elicits a significantly higher BOLD response than all other conditions (post-peak: Consistent (ß=-1.318 (SE.282), t=-4.677, p<.001), Inconsistent (ß=-.806 (SE=.282), t=.031, p=.004), Context Only (ß=.815 (SE=.282), t=2.891, p=.004), which is also found in the recovery time window (ps<.001). This effect is observed in V1 but does not reach significance for all conditions as we find here in V2.

We find no significant differences in between the Context only condition and either the Consistent or Inconsistent conditions in any of the three time windows (ps>.05). The BOLD response is similar in amplitude to trials in which the target region contains low-contrast information, and the surround contains high-contrast information, implying the BOLD response we observe is driven mainly by information in the surround, at least at the peak. Results are summarised in Figure 3.11.



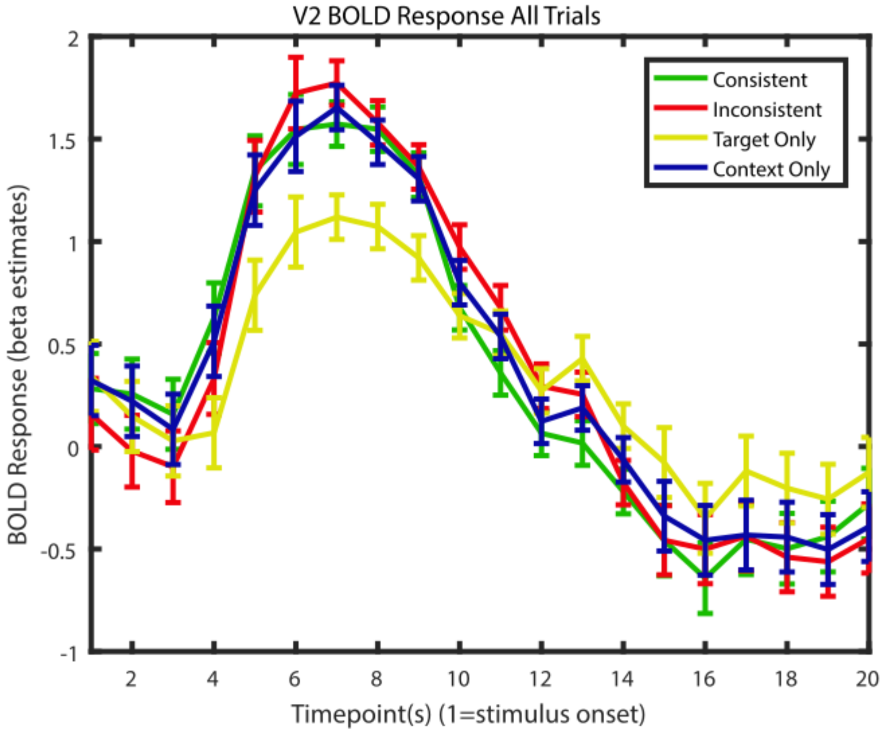


Figure 3.11 V2 BOLD response. Haemodynamic response of V2 target region ROIs across participants (N=22). Beta estimates (y-axis) are plotted as a function of time (seconds post-stimulus-onset, x-axis) for each consistency condition. The top figure includes the target mapping condition in which a high contrast stimulus is presented in the target region, driving a higher BOLD response. The bottom figure removes this condition to illustrate any differences between the low-contrast consistency conditions. Error bars represent Standard Error.

3.4.4.2 Correct Trials Only

Limiting our analysis to correct trials only, we see the same effects as with all trials being included (Figure 3.12). We find a significant difference between the Consistent and Inconsistent conditions in both the peak (ß-.787 (SE=.376), t=-.2903, p=.036) and post-peak (ß=-.766 (SE.376), t=-2.034, p=.042) time windows, with Inconsistent trials eliciting a significantly higher BOLD response amplitude than trials in which the surround and target region were Consistent. As with all trials, this response amplitude profile reverses in the recovery time window and the BOLD response amplitude for Consistent trials is higher than Inconsistent trials, however this difference does not reach significance (ß.322 (SE=.378), t=.851, p=.394).

BOLD response amplitude in the Target Only condition is lower than all other conditions at the peak, however this only reaches significance between the Inconsistent and Target Only condition (ß=1.239 (SE=.383), t=3.234, p=.001). The Context Only condition response is also lower at the peak than the two conditions which contain both surround and target information, however this does not reach significance (ps>.05).

In the post-peak and recovery time windows, we find a higher BOLD response in the conditions in which information is only available in either the target or surround region, compared to the conditions in which both types of input are present. The Target Only condition response amplitude is significantly higher than the Consistent condition (ß=-1.502 (SE=.358), t=-4.199, p<.001) and almost significant for the Inconsistent condition (ß-.736 (SE=.383), t=-1.920, p=.055). The Context Only condition is also only significantly higher than the Consistent condition (ß=-.864 (SE=.324), t=-2.667, p=.008) and not the Inconsistent condition (ß-.098 (SE=.353), t=-.279, p=.779). The Target Only condition elicits a higher response amplitude than the Context Only condition, but this does not quite reach significance in the post-peak time window (ß=.638 (SE=.333), t=1.916, p=.055).

In the recovery window, the distinction between Target Only and Context Only response amplitude increases, with the Target Only condition eliciting a significantly higher response than the Context Only condition (ß=.713 (SE=.335), t=2.126, p=.033). The Target Only condition response is also significantly higher than both the Consistent (β=-.938 (SE=.360), t=-2.609, p=.009) and Inconsistent (β=-1.260 (SE=.385), t=-3.275, p=.001) conditions whereas the response in the Context Only condition does not differ significantly from either Consistent or Inconsistent responses (ps>.05).

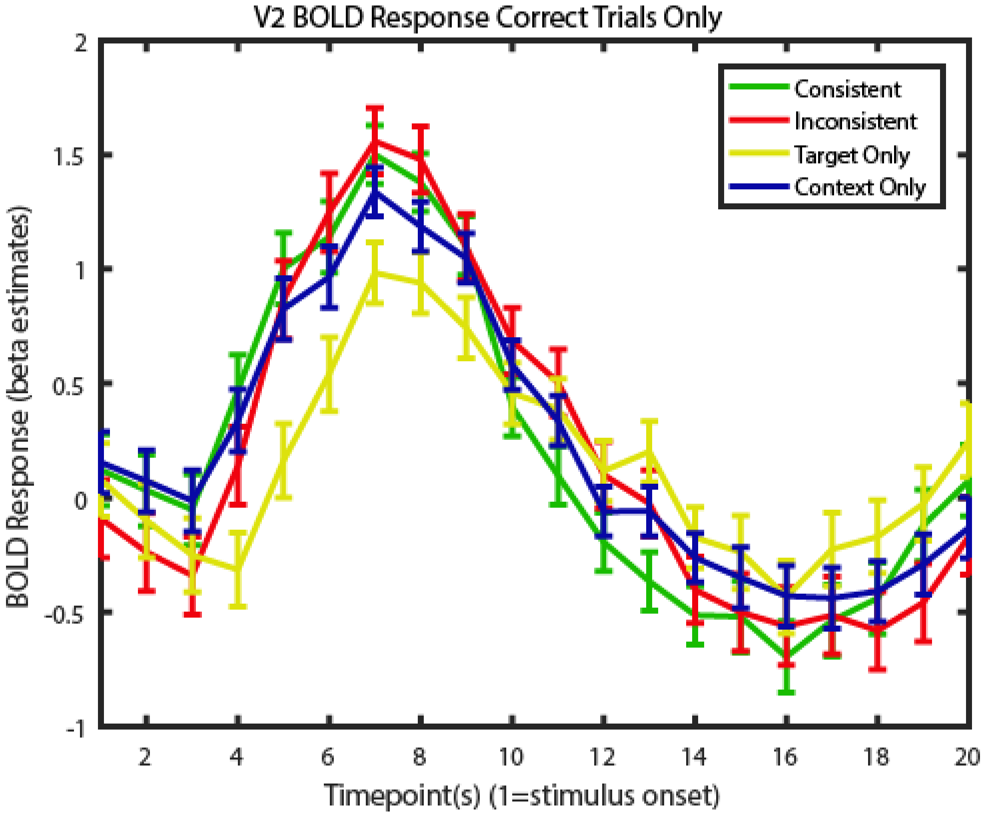


Figure 3.12 V2 BOLD response for correct trials only. Haemodynamic response for trials in which subjects responded correctly. Beta estimates are plotted as a function of time post-stimulus-onset for each consistency condition. Error bars represent Standard Error.

3.4.5 Multivariate Analysis V2

3.4.5.1 Cross-Validation

Multivariate cross-validation analysis on the V2 ROIs concerning the ability to decode the content of the target region reveals distinctions between some of the experimental conditions. The classifier can decode significantly above chance (p=.015) only in the Inconsistent condition (Mean=56.36%, Median=53.75%, SD=10.20%) but not in the Consistent (Mean=52.27%, Median=50%, SD=6.48%), Target Only (Mean=52.73%, Median=55.00%, SD=11.90%) and Context Only (Mean=52.50%, Median=52.50%, SD=10.72%) conditions which do not allow for decoding accuracy significantly above chance (p=.195, p=.286, p=.253, respectively). The mean decoding accuracies do not differ significantly between conditions (ps>.05). Figure 3.13 shows the cross-validation performance for each of the four conditions.

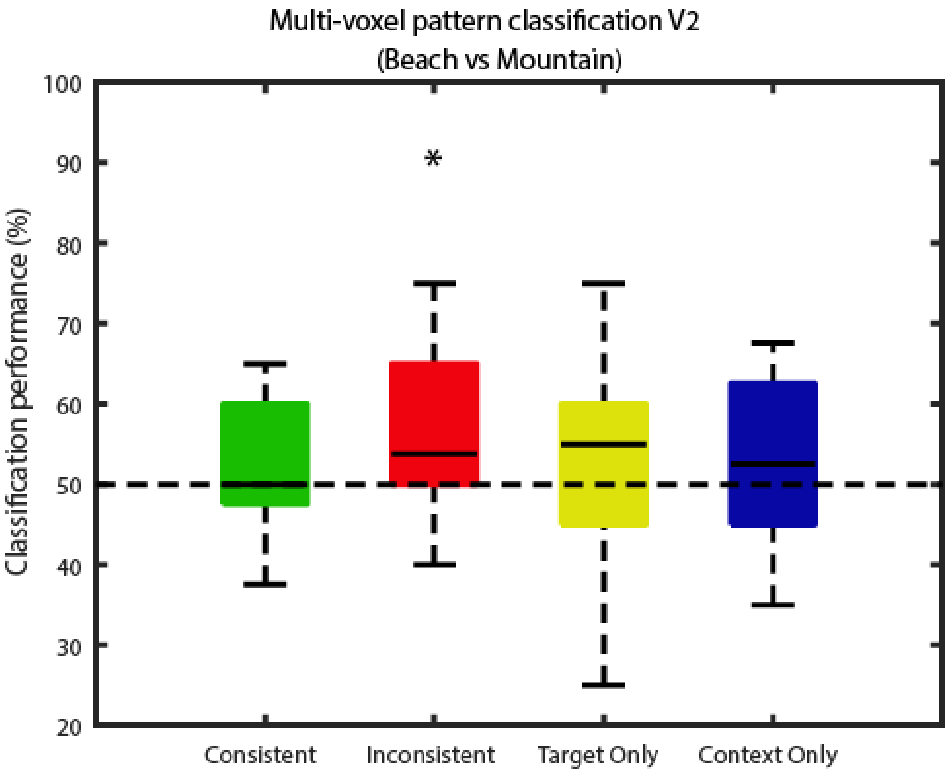


Figure 3.13 MVPA cross-validation results for V2. Multivariate cross-validation classification results for V2 ROIs (N=22). Classifier performance (%) is plotted for each consistency condition. The top and bottom edges of the box indicate the 75th and 25th percentiles, respectively, with the median denoted with a solid black line. The whiskers depict the range, excluding outliers (which are denoted by +). The dashed black line depicts chance-level performance (50%). Performance that differs significantly from chance is denoted by *.

3.4.5.2 Cross-Classification

We employed the same cross-classification analysis to V2 ROIs as we did to V1; we trained the classifier on purely feedback information (Context Only condition), purely feedforward information (Target Only) and consistent feedforward and feedback information (Consistent condition) in turn and tested on the remaining conditions.

When trained on purely feedback information, the classifier was not able to accurately decode above chance in any of the three conditions. Mean accuracy in the Consistent condition was 50.23% (Median=51.25%, SD=8.30) and this performance was not significantly above chance level (p=.741). Performance decreases in the Inconsistent condition with a Mean accuracy of 46.25% (Median=46.88%, SD=7.07%). This decreased performance accuracy is significantly below chance level (p=.016) implying that the contribution of incompatible feedforward signals interferes with classifier performance. When tested on the Target Only (i.e., purely feedforward information), performance again is around chance level (p=.346), (Mean=48.13%, Median=50.63%, SD=6.02%). Performance between conditions does not differ significantly from each other (ps>.05). Thus, it appears that when trained on feedback information from the surround, the classifier cannot accurately decode the content of the target region above chance level in V2 ROIs. Performance is hindered when incompatible feedforward information is present in the target region.

When trained on feedforward information the classifier was not able to accurately decode the content of the target region significantly above chance in V2 in the Consistent (Mean=52.39%, Median=41.88%, SD=5.60%, p=.050), Inconsistent (Mean=52.33%, Median=51.25%, SD=7.05%) or Context Only (Mean=48.13%, Median=50.63%, SD=6.02%) condition (ps>.05). Only the Inconsistent and Context Only conditions differed significantly from each other (p=.042).

When trained on the Consistent condition, the classifier is not able to decode significantly above chance in the Target Only condition (Mean=52.39, Median=51.88, SD=5.60). Performance in the Inconsistent (Mean=49.72, Median=48.75, SD=6.41) and Context Only (Mean=50.23%, Median=51.25%, SD=8.30%) are even closer to chance level (ps>.05). Therefore, when trained on contributions from both feedforward and feedback signals, the classifier cannot accurately decode the content of the target region. Performance between conditions does not differ significantly. Figure 3.14 depicts classifier performance in each of the three training and testing scenarios.

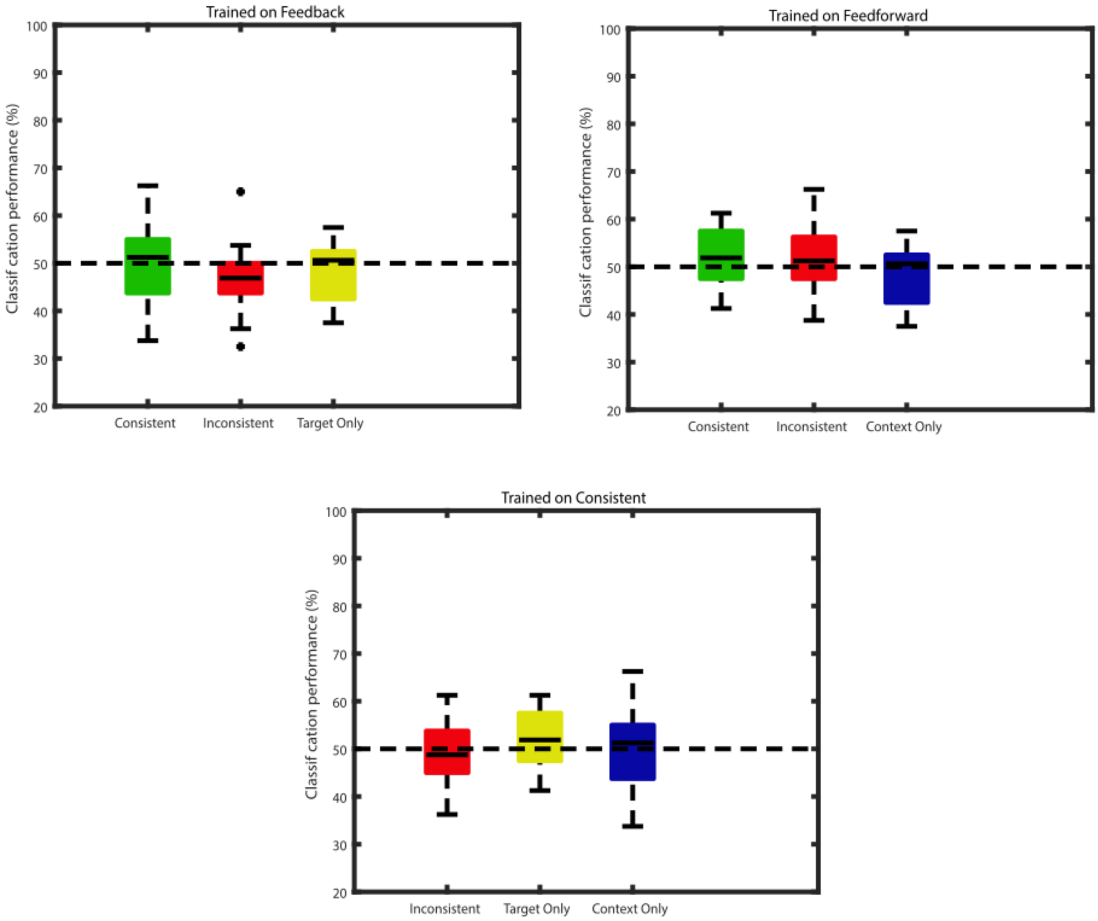


Figure 3.14 V2 MVPA cross-classification results. Multivariate cross-classification results for V2 target-region ROIs when trained on feedback (top), feedforward (middle) and consistent (bottom) information and tested on the remaining conditions. Classifier performance as a percentage is plotted for each consistency condition. The top and bottom edges of the box indicate the 75th and 25th percentiles, respectively, with the median denoted with a solid black line. The whiskers depict the range, excluding outliers (which are denoted by +). The dashed black line shows chance-level performance at 50%.

3.5 Results Higher Contrast Experiment

3.5.1 Psychophysics

In The Higher-Contrast Experiment, the low-contrast target region was shown at each subject’s 75% threshold contrast-level. I.e., the contrast level at which subjects detected a stimulus presence 75% of the time. Results of the psychophysical data in this experiment reveals the same pattern of results as previously observed. Consistent information in the surround increases performance accuracy (Mean=90.71%, SD=9.49) of the recognition task compared to trials in which there is no information in the surround (Mean=86.38%, SD=11.67), or the information in the surround is inconsistent (Mean=76.65%, SD=14.79). Compared to the lower-contrast experiment, performance accuracy is higher in each consistency condition. This reflects the higher contrast level of the target region in this experiment. We find a significant effect of consistency on performance accuracy (F(2, 46) = 11.346, p<.001). Bonferroni-corrected post-hoc tests reveal a significant difference between Consistent and Inconsistent trials and Inconsistent and Target Only trials but performance in the Consistent and Target Only trials is not significantly different. This implies that consistent contextual information does not necessarily improve perception of degraded information in the target region, (as performance is similar to instances in which there is no contextual information available). Instead, Inconsistent information in the surround hinders ability to identify low-contrast scene information. In other words, information in the context only influences recognition of low-contrast information if it is incongruent.

Analysis of the data using Signal Detection Theory reveals a slight overall tendency for subjects to respond according to the surround, (criterion value = -0.3586). Despite this, subjects overall show a degree of sensitivity in being able to determine the content of the target region (dPrime = 2.5088) which is above chance-level performance. Thus, subjects are reliably able to perform the task despite a slight tendency toward basing their decision on information in the surround. Figure 3.15 and Table 3.3 summarise the behavioural test results.

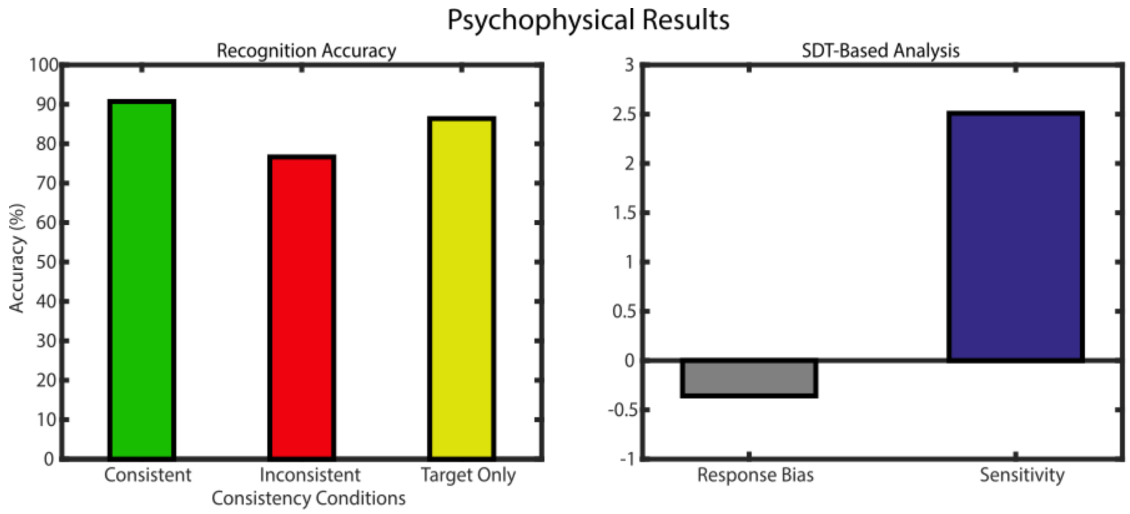


Figure 3.15 Psychophysical results for the higher-contrast experiment. The left figure displays performance accuracy as a percentage for each of the consistency conditions. The figure on the right shows the SDT-based measures of response bias (criterion) and sensitivity (dPrime).

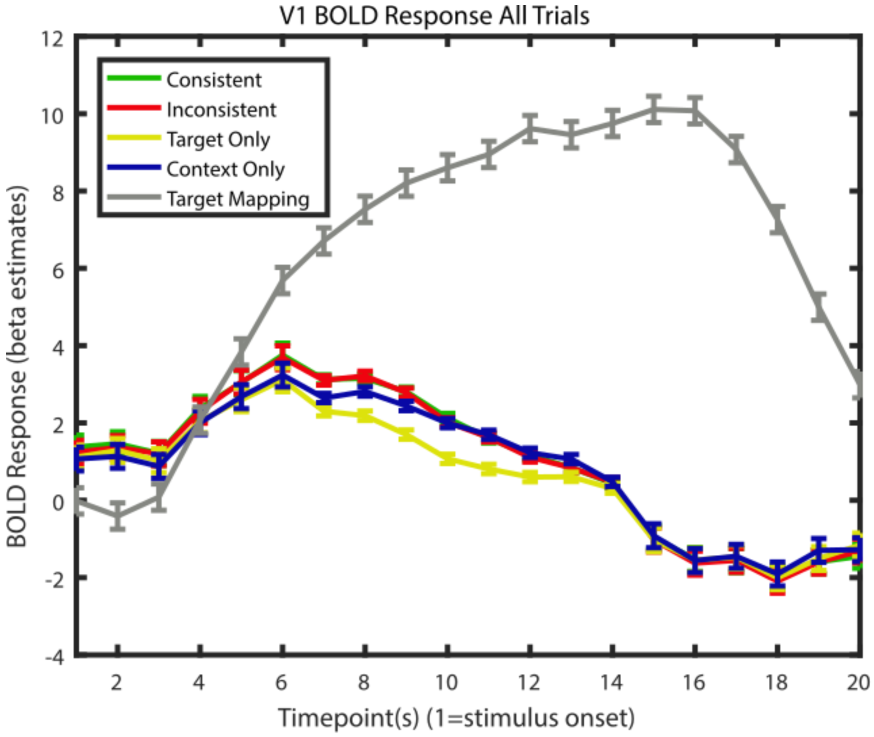
Hit Ratio	False Alarm Ratio	Response Bias (c)	Sensitivity (dPrime)
0.9071	0.2335	-0.3586	2.5088

Table 3.3 SDT results for the higher-contrast experiment. Results from the signal-detection-based analysis. The Hit ratio and False Alarm ratio are used to calculate response bias (criterion) and sensitivity (dPrime).

3.5.2 Univariate Analysis V1

3.5.2.1 All Trials

Looking at the BOLD response amplitude (Figure 3.16), we firstly note that there is a stronger BOLD response in this experiment compared to the lower contrast experiment, which would be expected due to the higher contrast of the target region.



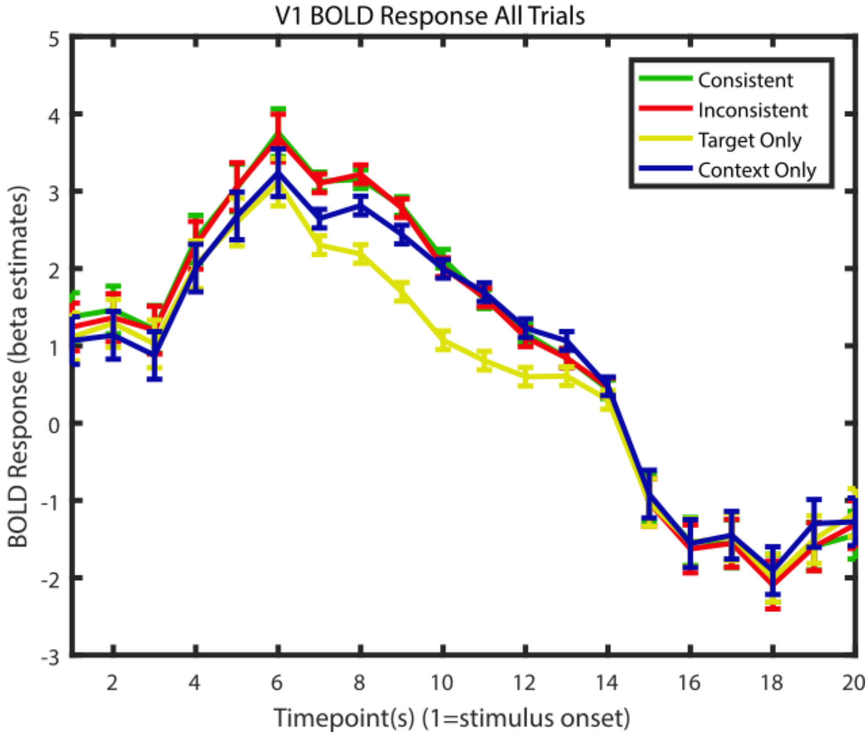


Figure 3.16 V1 BOLD response. Haemodynamic response in the V1 target region ROI across participants (N=24). Beta estimates are plotted as a function of time for each consistency condition. In the top figure, the target mapping condition is included in which a high contrast stimulus is presented in the target region, driving a higher BOLD response and providing perspective for the low-contrast conditions. In the bottom figure, this condition is removed and differences between the low-contrast target region conditions can be observed. Error bars represent Standard Error.

When looking for an effect of consistency, we find no significant differences between the Consistent and Inconsistent conditions in any of the three time windows (ps>.05). In the peak response window, the Context Only condition elicits a significantly lower BOLD response amplitude than both the Consistent (ß=.747 (SE=.300), t=2.489, p=.013) and Inconsistent (ß=.697 (SE=.300), t=2.322, p=.020) conditions. This distinction likely reflects the higher contrast of the target region in this experiment, meaning the BOLD signal is significantly reduced when there is no information in the Target region. As in The Lower-Contrast Experiment, the Target Only condition results in a significantly lower peak BOLD response amplitude than the Consistent, Inconsistent and Context Only (ps<.001) conditions. However, in this experiment, this reduction in response when contextual information is absent is less pronounced than in The Lower-Contrast Experiment in which the target region contained weaker feedforward information. Therefore, this finding likely reflects a stronger feedforward signal in the Target Only condition.

In the post-peak time window, the decreased BOLD response in the Target Only condition with respect to the conditions containing contextual information is maintained (ps<.05). The significant difference between the Context Only condition and the Consistent and Inconsistent conditions observed in the peak time-window disappears in the post-peak phase (ps>.05). The response amplitude is similar across all conditions with only a significant difference being observed between the Context Only and Target Only (ß=-1.369 (SE=.301), t=-4.551, p<.001) conditions however, this effect decreases as a function of time. The response amplitude for all conditions is similar in the recovery time-window, with the only significant difference being observed between the Inconsistent and Context Only condition (ß=-.661 (SE=.302), t=-2.188, p=.029). We do not observe the differences in undershoot profiles that we consider in The Lower-Contrast Experiment.

Again, it is worth noting for perspective the difference in signal amplitude between the experimental conditions and the target mapping condition, as with the lower contrast experiment, the target mapping condition elicits a much higher BOLD signal response amplitude. This is as expected as the target mapping condition stimulates the cortical target region with high contrast stimuli, compared to low-contrast scene information in the experimental conditions.

3.5.2.2 Correct Trials Only

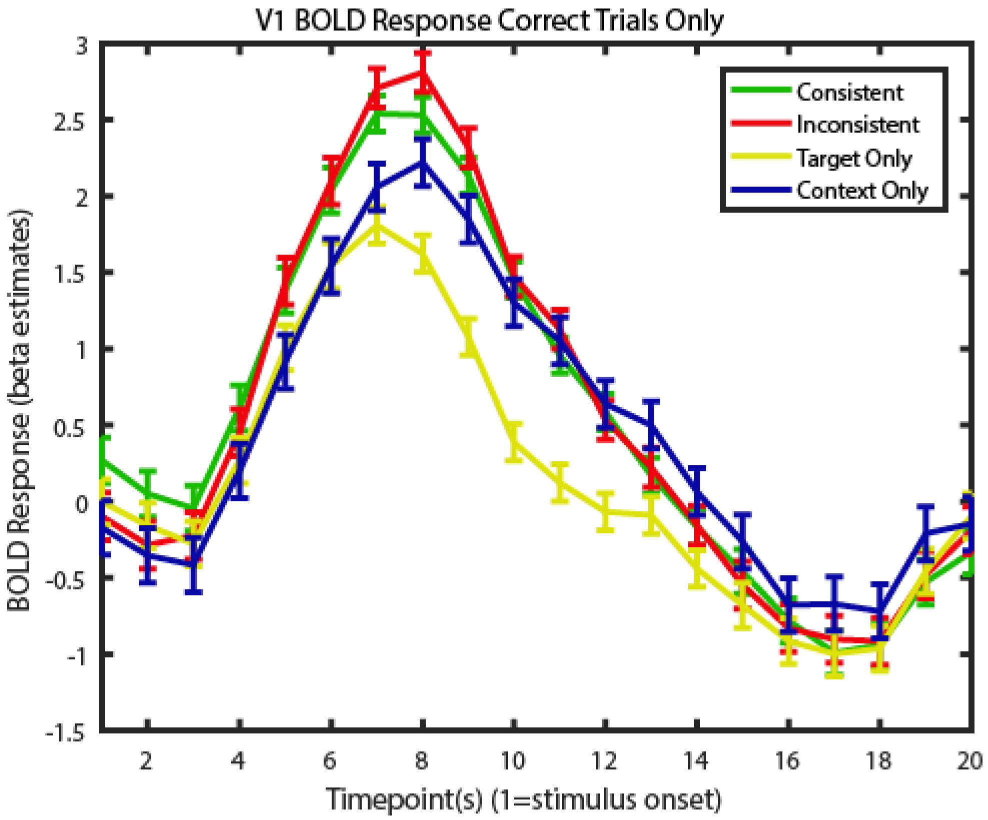


Figure 3.17 V1 BOLD response for correct trials only. Haemodynamic response function in the V1 target region for trials only in which the subjects responded correctly. Beta estimates are plotted as a function of time for each of the consistency conditions. Error bars represent Standard Error.

Focusing only on trials in which subjects responded correctly (Figure 3.17), we find an almost significant difference between the Consistent and Inconsistent conditions at the peak (ß=-.654 (SE=.335), t=-1.952, p=.051), with Inconsistent eliciting a higher BOLD response than the Consistent in correct trials. This difference does not persist beyond the peak and there is no significant difference between the two conditions in the post-peak or recovery time windows (ps>.05).

In the peak time-window, BOLD response in the Target Only condition is significantly lower than the Consistent (ß=3.862 (SE=.323), t=11.939, p<.001) Inconsistent (ß=4.516 (SE=.339), t=13.327, p<.001) and Context Only (ß =3.210

(SE=.380), t=-8.453, p<.001) conditions. The Context Only condition is also significantly lower than the Inconsistent (β=1.307 (SE=.390), t=3.353, p=0.001) but not the Consistent condition (p>.05).

In the post-peak time window, the significant difference between the Consistent and Inconsistent condition is eradicated (p=0.827) and both conditions lose response amplitude. The response amplitude in the Context Only condition decreases significantly less than the Consistent (β=-.812 (SE=.377), t=-2.153, p=0.031), Inconsistent (β=-.866 (SE=.390), t=-2.271, p=0.023) and Target Only (β=-2.215 (SE=.380), t=-5.824, p<.001) conditions, resulting in a higher response amplitude than conditions which contain feedforward information. The response in the Target Only condition remains lower than the other three conditions, significantly so for both Consistent (β=1.403 (SE=.324), t=4.330, p<.001) and Inconsistent (β=1.328 (SE=.339), t=3.914, p<.001) conditions.

In the recovery phase, we observe a less pronounced undershoot in the Context Only condition which is significant compared to the Consistent (β=-.961 (SE=.379), t=-2.535, p=.011), Inconsistent (β=-.849 (SE=.392), t=-2.165, p=0.030) and Target Only (β=-1.053 (SE.382), t=-2.756, p=.006) conditions. The other three conditions do not differ significantly from one another (ps>.05). It therefore appears that a lack of information in the target region results in less of an undershoot of the BOLD response amplitude.

3.5.3 Multivariate Analysis V1

3.5.3.1 Cross-Validation

In the cross-validation multivariate analysis (Figure 3.18), we find that the classifier is only able to perform significantly above chance in the Inconsistent (Mean=56.15%, Median=55.00%, SD=12.45%, p=.027) and Target Only condition (Mean=55.73%, Median=56.25%, SD=9.82%, p=.008) but does not perform above chance in either the Consistent (Mean=53.54%, Median=53.75%, SD=10.73%, p=.154) or Context Only (Mean=51.25%, Median=51.25%, SD=11.28%, p=.625). In other words, the classifier is only able to accurately distinguish scene information in the target region when the feedback signals are incompatible or absent in this experiment. Only the Target Only and Context Only conditions differ significantly from one another (p=.038), the other conditions have overall similar performance accuracies (ps>.05).

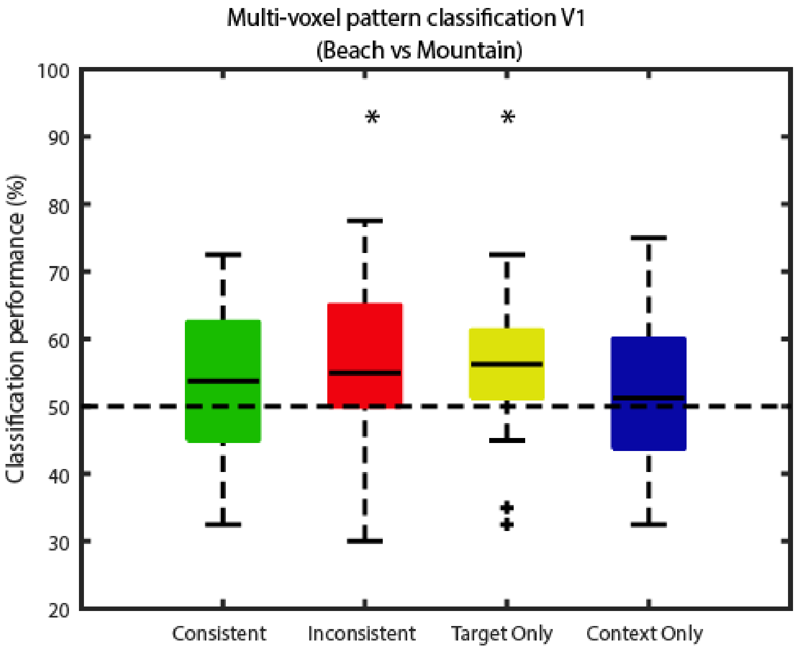


Figure 3.18 V1 cross-validation classification results. Classifier performance on the cross-validation analysis is plotted as a percentage for each of the four consistency conditions. Performance reflects the classifier's ability to decode a beach versus mountain scene in the target region. The top and bottom edges of the box indicate the 75th and 25th percentiles, respectively, with the median denoted with a solid black line. The whiskers depict the range, excluding outliers (which are denoted by +). The dashed black line depicts chance-level performance of 50%. Performance that differs significantly from chance is denoted by *.

3.5.3.2 Cross-Classification

In the first scenario, with the classifier trained on feedback information, we find that we are not able to decode significantly above chance in any of the three conditions (ps>.05). Despite this, performance is highest in the Consistent condition (Mean=52.45%, Median=50.63%, SD=7.05%) and drops significantly in the Inconsistent condition compared to Consistent performance (p=0.043), (Mean=46.72, Median=45.63%, SD=8.47%). The Target Only condition results in performance levels between the Consistent and Inconsistent conditions (Mean=50.16%, Median=47.5%, SD=6.97%).

Training on purely feedforward information results in chance level performance across all conditions (ps>.05). Consistent: Mean=50.47%, Median=50%, SD=5.76; Inconsistent: Mean=48.49%, 47.50%, SD=5.12%; Context Only: Mean=50.16, Median=47.50, SD=6.97%. Performance did not differ significantly between conditions (ps>.05).

Finally, when trained on consistent feedforward and feedback information, we also find that the classifier is unable to decode significantly better than chance in the Inconsistent (Mean=47.50%, Median=46.88%, SD=6.34%), Target Only (Mean=50.47%, Median=50.00%, SD=5.76%) or Context Only (Mean=52.45%, Median=50.63%, SD=7.05%), all ps>.05. We also find no significant differences between performance in any of the conditions (ps>.05). Classifier performance across the three cross-classification scenarios is summarised in Figure 3.19.

To sum, we don't find any evidence of significant cross-classification ability in target region V1 in any of the training/testing scenarios. We observe a general pattern in which decodability is reduced in the Inconsistent condition relative to the other conditions, but this is not enough of an effect to reach significance. The overall lack of findings with this cross-classification analysis could imply that when the feedforward signal is stronger (relative to The Lower-Contrast Experiment), the feedback signal patterns become less pronounced, influencing the classifier's ability to be able to distinguish the target region.

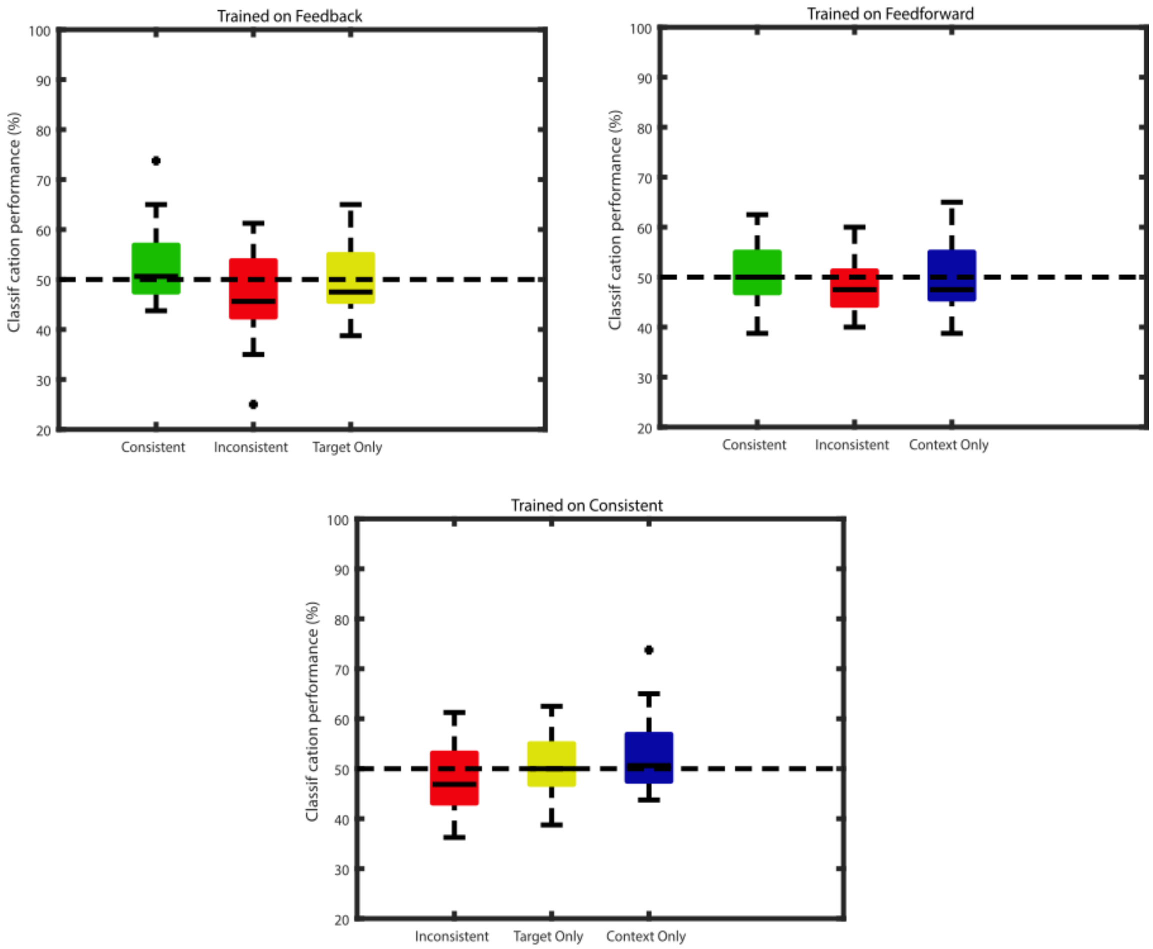


Figure 3.19 V1 cross-classification results. Classifier performance (as a percentage) plotted for each consistency condition for each of the three cross-classification scenarios. The classifier is trained on feedback (top), feedforward (middle) and consistent feedback and feedforward information (bottom) in turn and tested on each of the other consistency conditions. The top and bottom edges of the box indicate the 75th and 25th percentiles, respectively, with the median denoted with a solid black line. The whiskers depict the range, excluding outliers (which are denoted by +). The dashed black line represents chance-level performance at 50%.

3.5.4 Univariate Analysis V2

3.5.4.1 All Trials

Including all trials in a univariate analysis of BOLD response amplitude, the general response is lower in V2 than V1, an observation we also note in the lower-contrast experiment. BOLD response in this Experiment is higher than the V2 response in the lower-contrast experiment, likely due to the increased stimulus contrast in this experiment.

We find no significant differences between the Consistent and Inconsistent conditions in any of the three time windows (ps>.05). In fact, the response profile across all conditions is similar, an observation we do not find in The Lower-Contrast Experiment. Again, this can likely be contributed to the increased strength of the feedforward input.

In the peak time-window however, we still find a significantly lower BOLD response in the Target Only condition compared to both the Consistent (ß=.891 (SE=.291), t=3.064, p=.002) and Inconsistent (ß=1.115 (SE=.291), t=3.832, p<.001) conditions. The response between Target Only and Context Only is not significantly different. The Context Only condition response is significantly lower than the Inconsistent (ß=.768 (SE=.291), t=2.640, p=.008) but not the Consistent (p=.061) condition. It is worth noting that upon visually inspecting the haemodynamic response for V2 in this Experiment, the actual peak response appears to be earlier than we have accounted for in the previous sections (approximately 6s post-stimulus-onset compared with our previously defined peak of 8s post-onset), however, we have used the same time-windows to maintain consistency across analyses.

In the post-peak time-window we find a significantly lower response in the Target Only condition (versus the Consistent (ß=.994 (SE=.291), t=3.411, p<.001), Inconsistent (ß=.702 (SE=.291), t=2.410, p=.016) and Context Only (ß=-.814 (SE=.291), t=-2.792, p=.005) conditions). The Context Only condition response does not differ significantly from either the Consistent (p=.536) or Inconsistent (p=.705) conditions.

In the recovery timeframe, we observe a greater undershoot of the BOLD response in both conditions containing information in the surround and target regions. The Target Only and Context Only conditions have a similar response amplitude (p=.722). The Target Only condition differs significantly from the Inconsistent (ß=-.585 (SE=.293), t=-1.999, p=.045) but not the Consistent (p=.116) condition. The Context Only condition also differs significantly from the Inconsistent (ß=-.688 (SE=.293), t=-2.350, p=.019) condition and almost reaches significance when compared to the Consistent condition (p=.054). Figure 3.20 summarises the response profile across conditions.

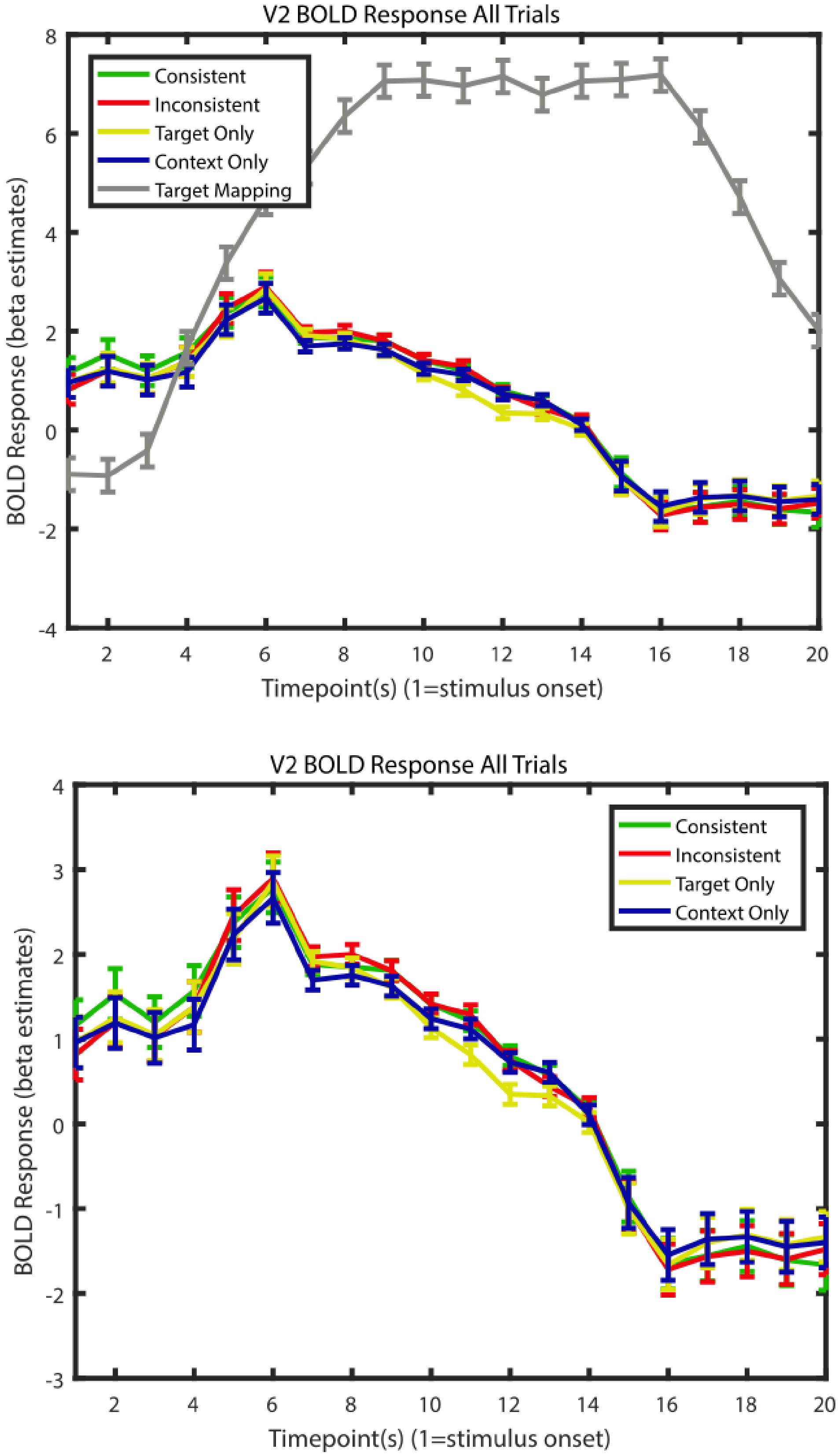


Figure 3.20 V2 BOLD response. Beta estimates as a function of time (post-stimulus-onset) for each consistency condition in V2 target region ROIs across subjects. The top plot includes the target mapping condition which stimulates the visual cortex with a high contrast flickering image, driving a higher BOLD response. The bottom plot removes this condition to closer inspect differences between the low-contrast consistency conditions. Error bars represent Standard Error.

3.5.4.2 Correct Trials Only

Focusing only on the trials in which the subjects responded correctly by eliminating any trials in which they responded incorrectly or did not respond at all, we find significant differences in the BOLD response amplitude in V2 (summarised in Figure 3.21).

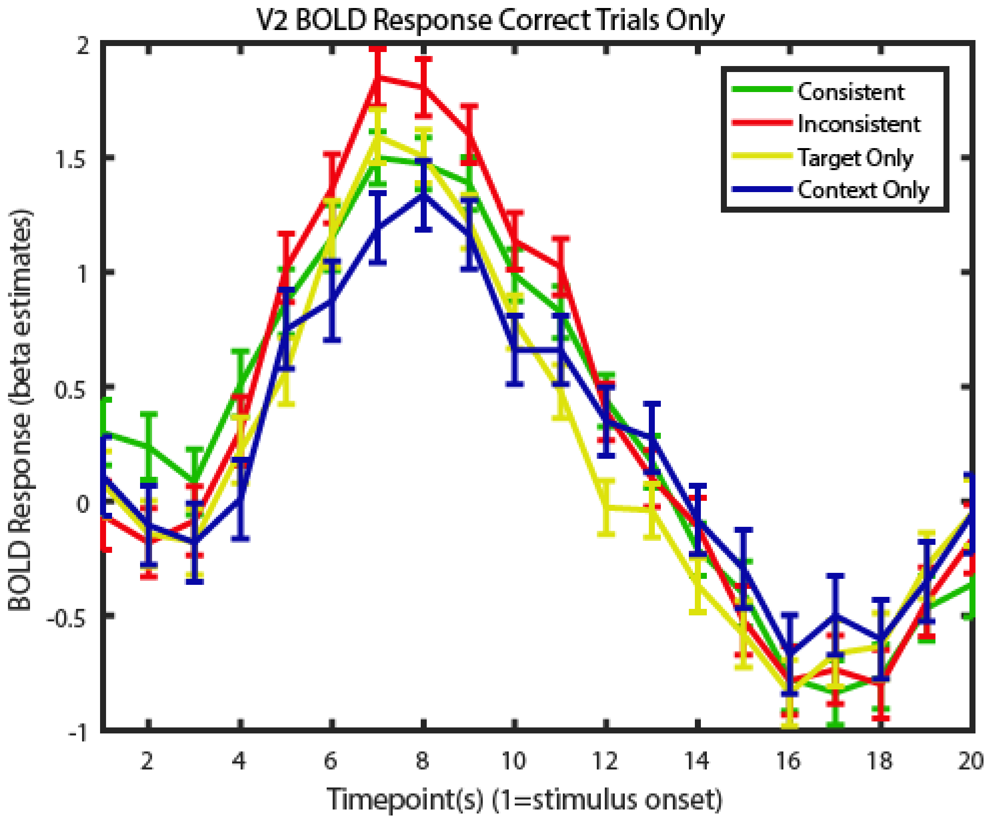


Figure 3.21 V2 BOLD response for correct trials only. BOLD response (as beta estimates) as a function of time post-stimulus-onset for each consistency condition in V2 ROIs for trials in which subjects responded correctly only. Error bars represent Standard Error.

In the peak time-window, the Inconsistent condition has a significantly higher BOLD response than all other conditions (Consistent: ß=-.892 (SE=.324), t=-2.757, p=.006; Target Only: ß=1.578(SE=.328), t=4.814, p<.001; Context Only: ß=1.746 (SE=.377), t=4.624, p=<.001). The Consistent condition also elicits a higher response than both the Target Only (ß=.686 (SE=.312), t=2.194, p=.028) and Context Only (ß=.853 (SE=.364), t=2.343, p=.019) conditions. The Target Only and Context Only condition responses do not differ significantly from each other (p=.648).

The distinction between the Consistent versus Inconsistent response decreases in the post-peak recovery window and does not reach significance (p=.668). The Target Only condition however has a significantly lower response than the Consistent (ß=1.012 (SE=.313), t=3.234, p=.001), Inconsistent (ß=.873 (SE=.328, t=2.661, p=0.008) and Context Only (ß=-1.261 (SE=.369), t=-3.421, p=.001) conditions. In contrast, the Context Only condition has a similar response to the Consistent (p=.496) and Inconsistent (p=.305) conditions.

In the recovery time window, the responses across conditions are similar, the Context Only has a slightly less pronounced undershoot than all other conditions but this does not reach significance (ps>.05).

Therefore, when we know subjects can correctly identify the content of the target region, there is a distinction in their BOLD response between conditions in V2, with Inconsistent trials eliciting a higher response.

3.5.5 Multivariate Analysis V2

3.5.5.1 Cross-Validation

Results of a cross-validation analysis (in which we train and test the classifier within each condition) reveal that the classifier is not able to accurately decode the content of the target region above chance in any of the four conditions (ps>.05). The Consistent condition resulted in the highest performance accuracy (Mean=52.76%, Median=55%, SD=9.71%), whereas performance was similar for the Inconsistent (Mean=50%, Median=50%, SD=9.90%), Target Only (Mean=50.92%, Median=50%, SD=9.94%) and Context Only (Mean=50.39%, Median=50%, SD=10.48%) conditions. The conditions did not differ significantly from each other (ps>.05). Figure 3.22 summarises the cross-validation classification results.

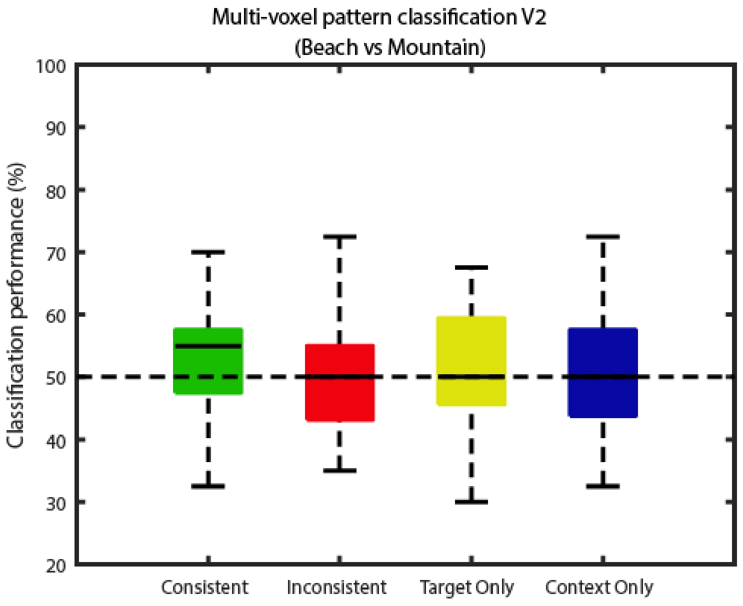


Figure 3.22 MVPA cross-validation results for V2. Classifier performance as a percentage is plotted for each consistency condition. The top and bottom edges of the box indicate the 75th and 25th percentiles, respectively, with the median denoted with a solid black line. The whiskers depict the range, excluding outliers (which are denoted by +). The dashed line represents chance-level performance at 50%.

3.5.5.2 Cross-Classification

When trained on purely feedback information (Context Only condition), we find a similar pattern of performance accuracy as we found in V1. The consistent condition resulted in the highest accuracy (Mean=51.64%, Median=52.50%, SD=6.90%). Performance dropped in the Inconsistent condition (Mean=48.95%, Median=47.50%, SD=6.85%) and was closest to chance in the Context Only condition (Mean=51.64%, Median=50%, SD=7.03%), however performance did not differ significantly from chance in any of the three conditions (ps>.05).

When trained purely on feedforward information (Target Only condition), the classifier again cannot decode significantly above chance in any of the three remaining conditions (ps>.05). In contrast to the scenario in which the classifier is trained on feedback information, when trained on feedforward information, testing on the Consistent condition resulted in the lowest performance accuracy (Mean=50.79%, Median=48.75%, SD=7.07%), compared to the Inconsistent (Mean=51.18%, Median=53.75%, SD=7.15%) and Context Only (Mean=51.64%, Median=50%, SD=7.03) conditions.

When trained on Consistent feedforward/feedback information, the classifier was unable to accurately decode the target region scene in the Inconsistent (Mean=51.91%, Median=52.50%, SD=7.43%), Target Only (Mean=50.79%, Median=48.75%, SD=7.07%) or Context Only (Mean=51.64%, Median=52.50%, SD=6.90%), (ps>.05). We also find no significant differences in accuracy between the conditions tested (ps>.05).

In sum, like classification performance in V1 in this Experiment, we are unable to distinguish between scenes of the target region using cross-classification. As we were able to do so in The Lower-Contrast Experiment when the feedforward information was weaker, the difficulty could be due to interference from the higher-contrast feedforward information present in this experiment. Figure 3.23 summarises the MVPA results in each scenario.

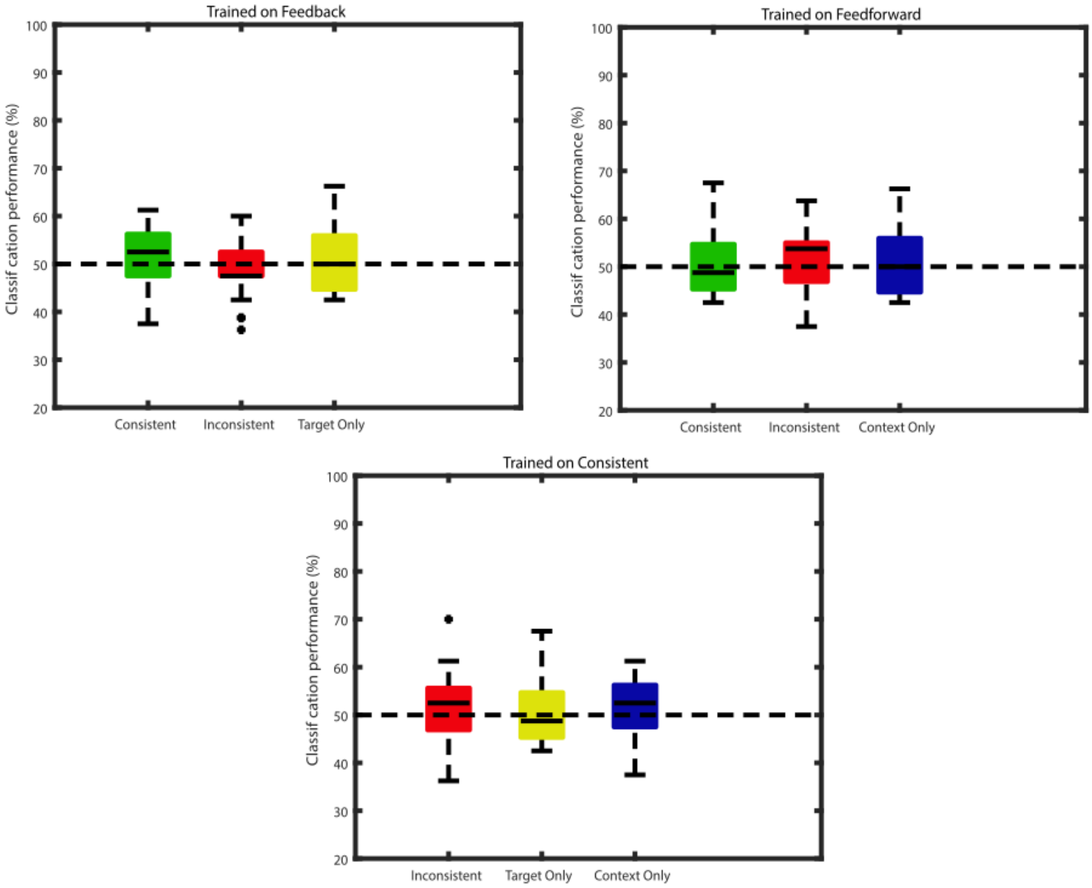


Figure 3.23 V2 MVPA cross-classification results. The figure depicts three training/testing scenarios in which the classifier is trained on feedback (top), feedforward (middle) and consistent feedback/feedforward (bottom) information in turn and tested on the remaining conditions. Classifier accuracy as a percentage is plotted for each consistency condition. The top and bottom edges of the box indicate the 75th and 25th percentiles, respectively, with the median denoted with a solid black line. The whiskers depict the range, excluding

outliers (which are denoted by +). The dashed black line depicts chance-level performance at 50%.

Summary statistics can be found below.

Lower-Contrast Experiment						
Region	**Time Window**	**Contrast**	**ß**	**SE**	**t**	**p**
V1	**Peak**	C vs I	0.301	0.282	1.067	.286
		C vs TO	3.784	0.282	13.416	<.001
		C vs CO	0.282	0.282	0.999	.317
		I vs TO	3.483	0.282	12.349	<.001
		I vs CO	-0.019	0.282	-0.067	.926
		TO vs CO	-3.502	0.282	-12.416	<.001
	Post-peak	C vs I	-0.181	0.283	-0.640	.524
		C vs TO	-0.790	0.283	-2.794	.005
		C vs CO	-0.640	0.282	-2.265	.023
		I vs TO	-0.610	0.283	-2.157	.031
		I vs CO	-0.459	0.283	-1.624	.104
		TO vs CO	0.151	0.283	0.533	0.596
	Recovery	C vs I	0.250	0.284	0.879	.379
		C vs TO	-1.062	0.285	-3.722	<.001
		C vs CO	-0.529	0.284	-1.865	.062
		I vs TO	-1.312	0.284	-4.618	<.001
		I vs CO	-0.778	0.284	-2.741	.006
		TO vs CO	0.534	0.284	1.876	.061
V1 Correct	**Peak**	C vs I	0.368	0.371	0.991	.322
		C vs TO	4.095	0.356	11.494	<.001
		C vs CO	0.702	0.322	2.180	.029
		I vs TO	4.208	0.379	11.093	<.001
		I vs CO	0.334	0.348	0.961	.337
		TO vs CO	-3.393	0.332	-10.230	<.001
	Post-peak	C vs I	-0.316	0.372	-0.850	.393
		C vs TO	-0.617	0.357	-1.728	.084
		C vs CO	-0.549	0.323	-1.700	.089
		I vs TO	-0.301	0.380	-0.792	.429
		I vs CO	-0.232	0.348	-0.667	.505
		TO vs CO	0.068	0.332	0.205	.844
	Recovery	C vs I	0.570	0.373	1.528	.126
		C vs TO	-0.929	0.359	-2.588	.010
		C vs CO	-0.436	0.325	-1.343	.179
		I vs TO	-1.499	0.381	-3.936	<.001
		I vs CO	-1.006	0.350	-2.873	.004
		TO vs CO	0.493	0.335	1.473	.141
V2	**Peak**	C vs I	-0.684	0.281	-2.435	.015
		C vs TO	0.724	0.281	2.577	.010
		C vs CO	-0.219	0.281	-0.777	.438
		I vs TO	1.409	0.281	5.011	<.001
		I vs CO	0.466	0.281	1.657	.097
		TO vs CO	-0.943	0.281	-3.354	<.001

	Post-peak	C vs I	-0.512	0.282	-1.820	.069
		C vs TO	-1.318	0.282	-4.677	<.001
		C vs CO	-0.504	0.281	-1.790	.073
		I vs TO	-0.806	0.282	-2.863	.004
		I vs CO	0.009	0.282	0.031	1.000
		TO vs CO	0.815	0.282	2.891	.004
	Recovery	C vs I	0.007	0.283	0.026	1.000
		C vs TO	-1.102	0.284	-3.873	<.001
		C vs CO	-0.198	0.283	-0.702	.484
		I vs TO	-1.109	0.283	-3.917	<.001
		I vs CO	-0.206	0.283	-0.726	.467
		TO vs CO	0.903	0.284	3.182	.001
V2 Correct	**Peak**	C vs I	-0.787	0.376	-2.093	.036
		C vs TO	0.452	0.357	1.266	.205
		C vs CO	-0.173	0.323	-0.535	.592
		I vs TO	1.239	0.383	3.234	.001
		I vs CO	0.614	0.352	1.743	.081
		TO vs CO	-0.625	0.332	-1.882	.060
	Post-peak	C vs I	-0.766	0.376	-2.034	.042
		C vs TO	-1.502	0.358	-4.199	<.001
		C vs CO	-0.864	0.324	-2.667	.008
		I vs TO	-0.736	0.383	-1.920	.055
		I vs CO	-0.098	0.353	-0.279	.779
		TO vs CO	0.638	0.333	1.916	.055
	Recovery	C vs I	0.322	0.378	0.851	.394
		C vs TO	-0.938	0.360	-2.609	.009
		C vs CO	-0.225	0.326	-0.690	.489
		I vs TO	-1.260	0.385	-3.275	.001
		I vs CO	-0.547	0.355	-1.540	.123
		TO vs CO	0.713	0.335	2.126	.033

Table 3.4 Statistics for the lower-contrast experiment. Summary statistics for each region (all trials and correct trials only) in each time window for the low-contrast experiment. Peak, post-peak and recovery time windows refer to 8-11s, 12-15s, and 16-19s post-stimulus onset, respectively. We ran the following 6 contrasts in each of these time-windows: C vs I – Consistent vs Inconsistent; C vs TO – Consistent vs Target Only; C vs CO – Consistent vs Context Only; I vs TO - Inconsistent vs Target Only; I vs CO - Inconsistent vs Context Only; TO vs CO - Target Only vs Context Only. We provide a beta-estimate which estimates the amplitude of the effect, followed by standard error of this effect and a t-value which is the associated statistic of the effect. Finally, we display the significance value (α=.05). Statistics are calculated through a deconvolution GLM-based analysis conducted on BrainVoyager software.

Higher-Contrast Experiment						
Region	**Time Window**	**Contrast**	**ß**	**SE**	**t**	**p**
V1	**Peak**	C vs I	0.050	0.300	0.167	.862
		C vs TO	3.924	0.300	13.069	<.001
		C vs CO	0.747	0.300	2.489	.013
		I vs TO	3.874	0.300	12.902	<.001
		I vs CO	0.697	0.300	2.322	.020
		TO vs CO	-3.177	0.300	-10.581	<.001

	Post-peak	C vs I	0.089	0.301	0.296	.769
		C vs TO	0.992	0.301	3.300	.001
		C vs CO	-0.376	0.301	-1.252	.210
		I vs TO	0.903	0.301	3.004	.003
		I vs CO	-0.465	0.301	-1.547	.121
		TO vs CO	-1.369	0.301	-4.551	<.001
	Recovery	C vs I	0.184	0.302	0.609	.545
		C vs TO	-0.164	0.302	-0.543	.585
		C vs CO	-0.477	0.302	-1.581	.114
		I vs TO	-0.348	0.302	-1.151	.249
		I vs CO	-0.661	0.302	-2.188	.029
		TO vs CO	-0.313	0.302	-1.037	.299
V1 Correct	**Peak**	C vs I	-0.654	0.335	-1.952	.051
		C vs TO	3.862	0.323	11.939	<.001
		C vs CO	0.652	0.376	1.734	.083
		I vs TO	4.516	0.339	13.327	<.001
		I vs CO	1.307	0.390	3.353	.001
		TO vs CO	-3.210	0.380	-8.453	<.001
	Post-peak	C vs I	0.075	0.336	0.222	.828
		C vs TO	1.403	0.324	4.330	<.001
		C vs CO	-0.812	0.377	-2.153	.031
		I vs TO	1.328	0.339	3.914	<.001
		I vs CO	-0.886	0.390	-2.271	.023
		TO vs CO	-2.215	0.380	-5.824	<.001
	Recovery	C vs I	-0.112	0.337	-0.332	.743
		C vs TO	0.092	0.325	0.283	.781
		C vs CO	-0.961	0.379	-2.535	.011
		I vs TO	0.204	0.341	0.598	.550
		I vs CO	-0.849	0.392	-2.165	.030
		TO vs CO	-1.053	0.382	-2.756	.006
V2	**Peak**	C vs I	-0.223	0.291	-0.768	.442
		C vs TO	0.891	0.291	3.064	.002
		C vs CO	0.545	0.291	1.872	.061
		I vs TO	1.115	0.291	3.832	<.001
		I vs CO	0.768	0.291	2.640	.008
		TO vs CO	-0.347	0.291	-1.192	.233
	Post-peak	C vs I	0.292	0.291	1.001	.316
		C vs TO	0.994	0.291	3.411	.001
		C vs CO	0.180	0.291	0.619	.536
		I vs TO	0.702	0.291	2.410	.016
		I vs CO	-0.111	0.291	-0.382	.705
		TO vs CO	-0.814	0.291	-2.792	.005
	Recovery	C vs I	0.125	0.292	0.427	.668
		C vs TO	-0.460	0.293	-1.572	.116
		C vs CO	-0.563	0.293	-1.924	.054
		I vs TO	-0.585	0.293	-1.999	.046
		I vs CO	-0.688	0.293	-2.350	.019
		TO vs CO	-0.103	0.293	-0.351	.722
V2 Correct	**Peak**	C vs I	-0.892	0.324	-2.757	.006
		C vs TO	0.686	0.312	2.194	.028
		C vs CO	0.853	0.364	2.343	.019

		I vs TO	1.578	0.328	4.814	<.001
		I vs CO	1.746	0.377	4.624	<.001
		TO vs CO	0.168	0.368	0.455	.648
	Post-peak	C vs I	0.139	0.324	0.428	.668
		C vs TO	1.012	0.313	3.234	.001
		C vs CO	-0.249	0.365	-0.681	.496
		I vs TO	0.873	0.328	2.661	.008
		I vs CO	-0.387	0.378	-1.024	.305
		TO vs CO	-1.261	0.369	-3.421	.001
	Recovery	C vs I	-0.085	0.325	-0.260	.799
		C vs TO	-0.424	0.315	-1.348	.178
		C vs CO	-0.720	0.367	-1.963	.050
		I vs TO	-0.339	0.329	-1.030	.304
		I vs CO	-0.636	0.380	-1.673	.094
		TO vs CO	-0.296	0.370	-0.800	.423

Table 3.5 Statistics for the higher-contrast experiment. Summary statistics for each region (all trials and correct trials only) in each time window for the low-contrast experiment. Peak, post-peak and recovery time windows refer to 8-11s, 12-15s, and 16-19s post-stimulus onset, respectively. We ran the following 6 contrasts in each of these time-windows: C vs I – Consistent vs Inconsistent; C vs TO – Consistent vs Target Only; C vs CO – Consistent vs Context Only; I vs TO - Inconsistent vs Target Only; I vs CO - Inconsistent vs Context Only; TO vs CO - Target Only vs Context Only. We provide a beta-estimate which estimates the amplitude of the effect, followed by standard error of this effect and a t-value which is the associated statistic of the effect. Finally, we display the significance value (α=.05). Statistics are calculated through a deconvolution GLM-based analysis conducted on BrainVoyager software.

3.6 Discussion

3.6.1 Outline of study

With this study we aimed to investigate how top-down predicted information influences the processing of dim feedforward input. To achieve this, we employed a partial occlusion paradigm in which a portion of the visual scene was shown at low contrast. The remaining full-contrast scene was either contextually consistent or inconsistent, meaning predictions formed on the basis of this context were either relevant or irrelevant. We looked at the psychophysical effects of consistency using a recognition task and also investigated the neural response to predicted or unpredicted low-contrast information at both the univariate and multivariate level.

In this discussion, I will look at what the results tell us about the individual role of feedback and feedforward input within this paradigm as well as their integration. I will then outline possible ways in which the two processes operate at the cortical level and how these results lend support to some important theories of cortical function and some emerging findings. I will outline some potential modulations and contributing factors to the interpretation of our findings. Finally, I will discuss potential caveats in the interpretation of these results and how they can be used to aid our understanding of how feedback and feedforward processing occurs in the brain.

3.6.2 Evidence of feedback signals in occluded cortex

We measured BOLD response amplitude in primary visual cortex as a function of time using a deconvolution analysis. Our univariate findings provide further support for the presence of feedback signals in occluded V1, in line with previous multivariate findings (Muckli et al., 2015; Smith & Muckli, 2010). We find a significantly higher BOLD response in occluded V1 in conditions in which there is contextual information present in the surround. When no contextual information is available, the BOLD signal is significantly reduced. These findings suggest that the neural response within the occluded target region of V1 is driven largely by feedback signals as their presence has a substantial influence on the amplitude of the BOLD response.

The evidence of feedback within this paradigm supports previous findings which reveal scene information from the surrounding context is present in unstimulated regions of visual cortex (Keller et al., 2020; Morgan et al., 2019; Muckli et al., 2015; Revina, 2021; Smith & Muckli, 2010; van Kemenade et al., 2020), a finding which extends to artificial neural networks (Ernst et al., 2019). These highlight a role for context in providing information about the visual scene when information is missing. Our results extend this finding to suggest that contextual information is still fed back to areas of V1 that contain degraded feedforward input. That is, in addition to unstimulated areas of primary visual cortex, information in the surround still largely drives a neural response when feedforward in formation is present.

Our multivariate findings further corroborate previous work. Cross-validation MVPA reveals we are able to decode information in the occluded region of V1 only when contextual information is present. As such, our findings are in line with that of Smith & Muckli, (2010) who were also able to decode scene information in occluded V1. Our results, however, suggest these feedback signals are still sufficient to decode contextual information even when weak feedforward information is present.

The presence of contextual information also appears to largely drive the BOLD response in V2, with a significant distinction between conditions which contain contextual information and those which don’t. The overall response in V2 is lower than V1, which could arise from the nature of the visual hierarchy and differences in retinotopic properties of the two regions. Smith and Muckli do not find much of a contextual effect in V2 however our V2 results largely mirror those found in V1 in terms of the presence or absence of context, indicating a modulatory effect of contextual information outside of the primary visual cortex. Earlier work on V2 has implicated its role within contextual processing to an extent (Bakin et al., 2000; Raizada & Grossberg, 2001).

3.6.3 Influence of degraded feedforward information

A unique feature of our experiment lies in the inclusion of degraded feedforward information in our occluded region. How much of the response signal can be attributed to this weak feedforward information alone? When no information is presented in the target region, (that is, the occluded region contains no feedforward information), the signal is reduced in both V1 and V2, albeit marginally, relative to the signal difference between context versus no context. This suggests feedforward information also contributes to the BOLD response due to the reduction in signal when this information is eliminated and is not simply overwritten by the stronger influence of feedback signals.

When feedforward input is stronger in the higher-contrast experiment, we find more of a distinction between conditions in which feedforward information is present versus absent than we do when feedforward input is weaker. This is expected if feedforward information is driving some of the BOLD response and therefore this provides evidence that this is indeed the case.

At the multivariate level, with threshold-level information in the target region, the target only condition is the only condition in which we cannot accurately decode significantly above chance. This implies the feedforward information alone is insufficient in being able to determine the content of the target region. Conversely, in V2, the Target Only condition is the only condition in which we are able to decode significantly above chance, suggesting that degraded feedforward information drives a BOLD effect in V2.

When the feedforward information is stronger in the higher-contrast experiment, we can only decode feedforward information in the target region, reflecting the relative contribution of this stronger feedforward signal. We can no longer decode feedforward information in V2, suggesting a potential modulation of ambiguity in V2 activation.

3.6.4 Interactions between the two processing pathways

The central aim of this study is to investigate the influence of top-down predictions on the processing of degraded feedforward information. We have thus far considered the separate contribution of both feedforward and feedback information by looking at the implications if either input is eliminated. However, the trials which contain both information channels permit us to investigate how these processing streams interact at the neuronal level.

While the presence of feedback information has a significant influence on the response amplitude in V1, the content of the feedback signals does not seem to have such an effect. Consistent and inconsistent signals from the surrounding context elicit similar response profiles in occluded V1, at least when feedforward information is severely degraded. When feedforward input is more perceptually visible, the congruency between feedforward and feedback signals appears to be more salient, as inconsistent information in the surround elicits a higher response in the occluded region of V1.

This notion is supported at the multivariate level by the results of a cross-validation analysis. We find that when the feedforward signal is more degraded, only the presence of contextual information influences the ability to decode scene information in occluded V1. The content of the contextual information does not appear to make a difference to the pattern of the BOLD response, with no significant differences being found between the consistent and inconsistent conditions. When feedforward input is stronger, we are only able to decode in the Inconsistent and Target Only conditions- supporting the notion from the univariate findings that an increased feedforward input contributes more to the BOLD response than when the feedforward input is weaker. We also find multivariate evidence suggesting inconsistency between the feedback and feedforward inputs is reflected in the neural response of V1. This could be due to the inconsistency being more easily detectable when the feedforward signal is stronger which manifests as a detectable discrepancy in the underlying BOLD response.

The contextual modulation of V2 activation appears to be more sensitive to the content of the feedback information. Inconsistent feedforward and feedback signals appear to drive a higher BOLD response in V2 than when the signals are compatible. This is observed even when the target information is shown at very low contrast. Thus, while both regions are influenced by the presence of contextual information, V2 may play a role in processing the content of this information. This is supported in part by findings of a cross-validation analysis in which the classifier is only able to decode the target region of V2 when the contextual information is inconsistent, at least when the feedforward input is weakest.

A cross-classification analysis is able to give us insight into the individual contributions of feedback and feedforward information to the cortical BOLD response. By training the classifier on one condition and testing on another, we can investigate how much of the signal between the two conditions is compatible and how much each signal type contributes to the overall response pattern. When trained on purely feedback signals (information from the surround only), the addition of feedforward information in the testing phase significantly influenced classifier performance bi-directionally. This pattern, although not always significant, is replicated across regions and within both Experiments. This indicates that although our findings so far have suggested that the type of information in the surround does not influence the neuronal response, here we find that when training purely on surround information, the content of the additional feedforward input has a modulatory effect on the content of the target region. In cross-validation while training and testing within the same conditions, we are able to decode in both the Consistent and Inconsistent condition to the same degree. We have taken this to mean that the presence of feedback information is important for being able to decode the content of the target region. However, cross validation cannot tell us to what extent feedback and feedforward information is contributing to the response pattern. Cross-classification results here therefore reveal that feedforward information does indeed contribute to the response pattern.

Cross classification using the feedforward information to train the classifier results in the inability to decode the target region in any of the other conditions. This could reflect a genuine lack of contribution of the feedforward signal to the overall response profile or could merely be due to the lack of strength of the feedforward signal in comparison to the full-contrast surround information. The distinction found between the Consistent and Inconsistent conditions when the classifier is trained purely on feedback information suggests the latter, as it reveals a contributing role of the content of the feedforward signal, implying that the feedforward signal does play a role in the overall response pattern but is not sufficient on its own to enable accurate decoding.

Finally, cross-classification in which the classifier is trained on the Consistent condition (i.e., compatible feedforward and feedback signals) and tested on the remaining conditions reveals we are only able to accurately decode in the Context Only condition in V1. This again reinforces the driving role of feedback information in the overall response pattern in PVC. Testing on the Inconsistent condition decreased performance, highlighting the contribution of the feedforward input. This effect is more prominent in the low-contrast Experiment, implying that feedback signals have a stronger influence with a more degraded feedforward input. It is also only found in V1 as differences between conditions in V2 are non-significant and decoding in general is not above chance level.

Taken together, the results allude to a segregation of feedforward and feedback information within the response pattern in early visual cortex (a notion recently reported by Semedo and colleagues (Semedo et al., 2022)). Cross-classification reveals distinctions between the effects of feedforward and feedback contributions meaning signals from both processing streams are able to be teased apart within the BOLD response. This implies that contextual information from the surround has a modulatory effect on the processing of the feedforward input, but this may not be as straightforward as a simple combination of the signals, at least at the stage of early visual processing. It seems that this contextual effect is a result of feedback signals that remain distinct in their representation. This could be tested statistically by comparing the linear combination of both feedforward (only) and feedback (only) modulatory effects with the actual signal in trials in which both feedback and feedforward information is present.

A noteworthy finding lies in the recovery time-window of the haemodynamic response. Here, we find a cross-over of BOLD response amplitude between conditions which contain one type of information (either feedforward or feedback) versus conditions which contain both types of information. Generally, we find that the presence of both feedback and feedforward information results in a larger undershoot of the haemodynamic response. However, when the feedforward input increases in The Higher-Contrast Experiment, the undershoot response is similar for feedforward information only (Target Only) as it is for a combination of both types of signals. It appears that the more information available, the more of an undershoot of the BOLD response we observe. This is in line with findings from (Sadaghiani et al., 2009) who find a stronger post-stimulus undershoot for stronger feedforward input. They propose this undershoot effect reflects underlying neuronal activity rather than simply a vascular consequence as the effect persists even when vascular contributions are controlled for. In the context of our experiment, the undershoot phenomenon simply serves to reiterate the components of the response signal as being comprised of both feedback and feedforward information, as opposed to revealing any potential underlying contextual differences.

3.6.5 What are the implications?

We were not only interested in the neuronal responses to contextual modulation, but we also tested the behavioural implications by employing a recognition task. Subjects were asked to identify the content of the target region with every stimulus presentation. Results of the psychophysical task reveal performance was significantly reduced in terms of accuracy when the surround and target regions were inconsistent. This could imply that rather than an enhancement of recognition accuracy when relevant contextual information is present, irrelevant context in fact hinders performance, suggesting strong feedback signals that are incompatible with the weak feedforward information hamper subjects' ability to determine the content of the low-contrast region at the behavioural level. However, statistical analysis comparing Inconsistent trial performance to chance-level performance would be needed to confirm this effect.

When a response bias toward the information in the high-contrast surround is controlled for, we find that subjects are able to accurately determine the content of the target region (evident through sensitivity levels above chance), however, the nature of this analysis does not permit this to be tested within each consistency condition, so we cannot determine the relative sensitivities for each condition when a response bias is accounted for. Thus, performance accuracy implies that contextual consistency influences behavioural performance, however, we cannot rule out the possibility that this effect is at least in part, driven by a behavioural bias towards responding according to the information in the surround.

Neuronally, we find robust evidence (through both univariate and multivariate analyses) of the existence of feedback signals in occluded areas of visual cortex. The role of feedback is implicated within a wealth of findings across many sensory modalities and using a multitude of different neuroscientific techniques, and its importance in visual processing is undisputed. The exact nature of how concurrent processing occurs in the cortex however is still yet to be fully understood. Several theories and frameworks have encompassed the interaction between feedforward and feedback signals as a central hypothesis. Apical amplification (Phillips et al., 2016; Phillips, 2017) for example, proposes that contextual information in the form of feedback signals can influence how the cell responds to feedforward information. Specifically, when the top-down and bottom-up signals are compatible, the output of the cell is amplified. At the neural level, feedback information targeting the apical tufts of the layer 5 pyramidal cells can actively influence the action potentials generated in response to feedforward input to the soma. A mechanism in which this can operate is via BAC propagation (Larkum, 2013) in which compatible signals arriving at the active dendrites results in a burst of action potentials from the cell body.

Apical amplification is a candidate cortical operation that could be in play in situations in which feedforward input is weak, for example if it is degraded in some way or the input is ambiguous. If the weak feedforward signal is accompanied by compatible contextual information, the signal could be amplified. Incompatible signals on the other hand would be dis-amplified. This does not however, harmonise with our findings of the BOLD response. We largely find that the content of the contextual information does not impact the magnitude of the BOLD response, so although feedback signals are clearly crucial in driving the response, the information they contain does not appear significant at this level of processing. In fact, we find some indication that there is an enhancement at the cortical level of inconsistent information. This is more in line with predictive coding (Rao & Ballard, 1999) accounts of cortical processing which propose that the brain operates on the principle of minimising error (Friston, 2005) and as such error signals are deemed more salient and it is those signals which are allocated neural resources. The proposal by Bar and colleagues (Bar, 2007) in which there is a rapid activation of top-down information from bottom-up signal also implies an integration of both information streams within visual processing, in line with Rao & Ballard and Friston's accounts of predictive processing, and indeed each can be interpreted as highlighting different aspects of the same process. Bar's proposal would suggest bottom-up input initiates analogies and associations stored within memory which form predictions as to the interpretation of the bottom-up signal. The predictions are then compared to the bottom-up input and discrepancies are represented as an error signal according to Rao & Ballard and any consistencies are 'explained away' in order to minimise free energy, in line with Friston's suggestions. Thus, it is possible that the three proposals interplay during visual processing.

There are several potential explanations for our findings. Firstly, there could be multiple mechanisms and processes in play at once, which could account for different aspects of the relationship between contextual information and feedforward input. Perhaps different processes are used under different conditions, such as the level of ambiguity or the level of discrepancy between the context and input. If this is the case, it is likely these would be difficult to tease apart without actively setting out to do so a-priori and designing the experiment accordingly. For example, within subjects, we could parametrically manipulate the level of ambiguity and thus the reliance on feedback information. This could be achieved by altering the strength of the feedforward signal (meaning more emphasis is required on top-down information) within the one experiment. Alternatively, the strength of the feedback signals could be manipulated to test for bottom-up processing capabilities. Differences in discrepancies between the two information streams could be tested by including more levels of consistency or introducing different types of consistencies (e.g., temporal or spatial). The number of manipulations achievable within one experiment however is limited in fMRI in the interest of time and comfort for the participant when more predictors are included, and the number of trial repetitions required for multivariate-level analysis.

Secondly, we set out to investigate how top-down and bottom-up signals interact in early visual cortex when feedforward input is degraded but perhaps the integration of such signals is not specifically a process of primary visual cortex. Levelt and colleagues recently observed segregation of feedback and feedforward information in rodent V1 using two-photon calcium imaging within a partial occlusion paradigm. Neurons which did not respond (or showed a suppressed response) to fully visible natural scenes, showed a strong response to occluded images, suggesting a neural distinction between feedforward and feedback information (findings revealed through recent communication with Muckli & Petro, 2021). With cross-classification performance in our experiment not being highest when feedforward and feedback signals are consistent, we may too find evidence of segregation of these two inputs in primary visual cortex. While the presence of contextual information influences the response in PVC, the content of these signals may only be required at a higher processing stage.

3.6.6 Modulatory effects

Through our experimental design, we capture two potential regulatory factors on the contextual influence of early visual cortex, strength of feedforward input and level of hierarchical processing. Looking at differences between the two experiments, (i.e., comparing different strengths of feedforward information), we find a generally stronger response in both regions when the feedforward information is presented at higher contrast (although this is not tested for significance). We also find less of a distinction between conditions where both types of information are presented versus those in which one information source is eliminated. This implies that higher-contrast feedforward information contributes more to the overall response. In other words, when the feedforward signal is further degraded, the neural response is driven more by feedback information from the surround. When more feedforward information is available, the response in visual cortex is driven by both types of signals. We do not however, find a modulatory effect of the strength of signal on contextual modulation. It was hypothesised a-priori that increasing the contrast of the target region in the second experiment may enhance neuronal disparities between the content of the feedforward and feedback signals, however, we generally do not find this to be the case.

We explored two main regions of interest in early visual cortex to investigate our effect: V1 and V2. Both regions are part of the early processing stage of the visual system but both have been shown to perform distinct and often discreet operations (Goebel et al., 2012; Grill-Spector & Malach, 2004b). The main distinction we find in our results is that feedback information tends to drive the response in V1, with the BOLD signal being dependent on the presence of context. In V2 however, we begin to observe an implication of the content of the feedback signal and also find a driving effect of feedforward information in this region.

Through our findings we have identified two further factors which could modulate the effect of context on the processing of degraded input. The first factor is attention. In The Lower-Contrast Experiment, we initially tested participants (N=7) using a fixation cross which did not change for the duration of each trial. We then considered that in the Target Only condition (when only very low-contrast information in the target region was presented and no information was present in the surround), subjects may not recognise that a trial has begun. The target region stimulus is presented at perceptual threshold and therefore subjects will not accurately be able to recognise that a trial has been presented without the information in the surround being present. On subsequent subjects (N=18), we altered the fixation cross to change colour when the stimulus image was presented, thus indicating when a low-contrast scene was present in the Target Only condition. Comparing responses to the Target Only condition of the two groups of subjects, we observe a distinct difference between the two. When subjects were potentially unaware of the stimulus presentation, we see a much lower BOLD response to the Target Only condition than subjects who were made aware of the trial by a fixation cross colour change. This may provide evidence of an attentional effect driving the BOLD response. Engaged attention in task demands results in a markedly higher response amplitude than if attention is not necessarily engaged, implicating higher-level processing in the response profile of primary visual cortex during contextual modulation. This, however, would need to be tested for significance before any firmer claims can be made.

The second factor is task performance. In The Higher-Contrast Experiment, the peak response in V2 appears to be earlier than V1 in both experiments and earlier than V2 in The Lower-Contrast Experiment. Generally, the peak is found around 8s post-stimulus onset, whereas in V2 of The Higher-Contrast Experiment, the peak appears to be around 6s post-onset. However, when we focus only on the correct trials, the peak appears to match that found previously. This observation, in addition to the inconsistency effect in V2 only becoming prominent in the correct trials in this experiment (but still being present) suggests task performance in the higher-contrast experiment may confound the neuronal responses. The task is (at least in theory) easier in the higher-contrast experiment as the target region is more visible, therefore perhaps the trials which subjects respond incorrectly or don't respond to are more of an indication that they did not follow the task on these trials. In the first experiment, the target region is shown around threshold level contrast and incorrect/missed trials could instead reflect task difficulty (an effect explored by (Chen et al., 2009)) or perceptual threshold rather than a lack of concentration which could explain why isolating correct trials in the first experiment does not greatly influence the effects we observe. The nature of the task itself could also have a confounding effect on the neuronal response. In such an explicit recognition task like ours, the dependency on feedback signals may differ to a less explicit visual task such as a one back repetition task (used by Smith & Muckli (Smith & Muckli, 2010)), perhaps exploring different task types within the same paradigm would reveal any effects of task or task performance on neural activity during occlusion.

3.6.7 Caveats

Although we find strong evidence to support the prevalence of feedback signals in early visual areas, how can we determine that these signals do not instead reflect input from lateral connections? Feedforward connections to V1 are outnumbered not only by feedback, but also lateral connections, which are involved in processes such as perceptual grouping and contour integration (Gilbert & Li, 2013). It is possible that within this paradigm, activation found in unstimulated cortex is projected from stimulated cortex via lateral connections. We cannot rule the potential impact of these horizontal connections out completely however, Smith & Muckli controlled for such input via a weighted classification analysis and found that these connections are unlikely to be solely driving the activity found in the occluded region. As we employ a very similar paradigm, we can presume to a degree that the same would be found for our results. Additionally, if activity in our unstimulated ROIs is being driven by lateral connections, we may expect to find more of a distinction pertaining to the content of the feedforward information due to their role in contour integration. Gilbert & Sigman (Gilbert & Sigman, 2007) outline a potential role of feedback connections in gating horizontal connections which could highlight a complex interplay between three information streams.

Another critical point could be that the activity we observe in occluded cortex could be a result of spill over effects from the surround region. Extra-classical receptive field properties of primary visual neurons mean that visual stimulation in the surround region using high contrast scene images could be activating the receptive fields of some neurons within the target region (Angelucci et al., 2002), particularly near the border between the surround and target. It is difficult to know for certain that this is not the case without mapping the population receptive fields and analysing only voxels with PRFs in the target region, which we did not do in this experiment. However, we used mapping blocks to identify the ROIs as voxels responding to the target region while eliminating those responding to the border and surround regions using a conjunction contrast. This therefore eliminated voxels pertaining to the border between the surround and target regions, allowing a buffer that accounts for surround contamination. Thus, our selection of ROI voxels was as conservative as we could make it to avoid any spill over effects from the surround and we are fairly confident that the target region activity we observe is due to feedback information as a result. Smith & Muckli also rule out any significant influence of spill over using their weighted classification analysis, as they do not find differences in weighted classifier performance near the border between the surround and occluded region.

Our lack of definitive results could also reflect limitations of the study design. For example, the stimulus contrast was tailored for each subject's threshold detection level to ensure the influence of individual differences in visual perception were controlled for during stimulus presentation. However, we based this threshold on subjects' ability to detect a stimulus in the target region. Results from our experimental series in Chapter 2 suggest that contextual information does not influence detection rate. Therefore, it may have been more appropriate to use subject's recognition threshold to generate the stimuli given the fact our experiment here used a recognition task. There may be individual differences in ability to detect versus recognise an object, with some subjects having similar thresholds for each task, whereas others may have markedly different detection and recognition thresholds. Creating the experimental stimuli using subjects' recognition threshold would eliminate a confounding effect of this variability which we do not account for in this study. One final caveat to highlight is the indirect relationship between BOLD signal and neural activity (Logothetis, 2008). Instead of measuring the neural response directly, BOLD is a measure of the metabolic consequences of such activity and therefore any inferences drawn rely on the assumption that BOLD signal reflects underlying neural activity. Nevertheless, it is generally deemed a reliable measure of cortical signalling due to multimethod studies and coincidental findings between different methodologies.

3.6.8 Conclusion

To conclude, our findings reveal neuronal evidence of feedback signalling within early visual cortex, when degraded feedforward input is present. The presence of such signals drives both univariate and multivariate response patterns revealed through fMRI. It is less clear *how* these signals interact to ultimately influence the processing of weak feedforward input as we do not find a distinct difference between predictable and unpredictable top-down effects. Perhaps the findings are limited by the spatial resolution at 3T and higher-resolution investigation at the columnar or laminar level could provide further insight.

4 Ultra-high field fMRI – Review and project proposal

4.1 Abstract

This chapter constitutes the final PhD project, an investigation into the layered function of the cortex. Originally this was planned to be executed as a 7T fMRI project aiming to look at the role of feedback within the different layers. However, the circumstances of 2020 resulted in a change of plan and now the chapter will be presented as a literature review of ultra-high resolution fMRI followed by a project proposal for a 7T study. Despite this project not going ahead as planned, it is still important to review the literature to gain an understanding of ultra-high field MRI and its application. It is a recent advancement in neuroscientific methodology and thus even more crucial to understand the way in which it is being utilised and any potential problems it may bring. Without any background understanding, high-resolution and ultra-high-resolution fMRI may seem similar, however advancements such as this open up many opportunities as well as many issues to overcome and therefore a breadth of knowledge than encompasses this technique is vital for someone who wishes to continue in their research career.

4.2 The layered cortex

The mammalian cortex is divided into six distinct layers. These were first identified in the mid- 19th century by Ballanger (1840) who noticed that (to the naked eye) the cortical architecture was striped in appearance. This notion was reinforced by early staining and microscope work by Berlin in 1858, who noted differences in cell body architecture across the six layers. Since these pioneering observations, the six-layered cortex has been extensively investigated using a number of neuroscientific techniques. This has allowed for a relatively well-established understanding of the structural and functional connectivity profile of each layer.

As in the rudimentary work defining cortical architecture, the cortical layers have been identified over time using their differences in cell prevalence and arrangement as well as cell body and myelination density (Palomero-Gallagher &

Zilles, 2019). Laminar differences in dendritic and axonal properties have also been observed (Mohan et al., 2015). The six cortical layers are divided into nine neuronal layers (1, 2, 3, 4a, 4b, 4cα, 4cß, 5 and 6), with layer 1 on the outer surface, closest to the pial and layer 6 on the inner surface closest to the white matter (Goebel et al., 2012). These differences ultimately play a role in the functional distinctions between the laminae (Trampel et al., 2019). Using invasive techniques mainly in non-human primates, Lund and colleagues (Blasdel & Lund, 1983; Hawken et al., 1988; Rockland & Lund, 1983) outlined properties of each layer in primary visual cortex and their relative connections. See Figure 4.1 for an illustration of the cortical layers.

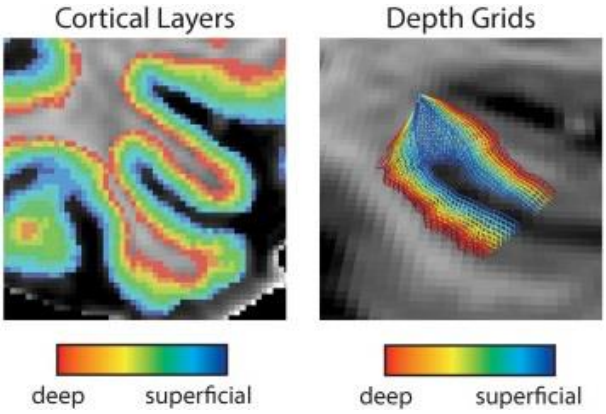


Figure 4.1 Illustration of cortical layers. Depiction of the cortical layers on the occipital cortex acquired with ultra-high resolution (9.4T, 0.35mm isotropic) fMRI. 2D depiction of the layers is on the left which can be used to create 3D cortical grids which define the layers across depths. Image obtained and adapted with permission from (De Martino et al., 2018)**.**

In primary visual cortex, Layer 1 is comprised mainly of dendritic and axonal connections, receiving inputs from the LGN, thalamic nuclei, extrastriate areas as well as subcortical regions. This layer is sparse in neural cell bodies but is salient in providing network connections with several stages of the cortical hierarchy. Layers 2 and 3 are known as the supragranular layers and consist of pyramidal cell bodies and dendrites. They receive input from layer 4 subdivision c and also connect with layer 5 neurons. They project to extrastriate cortex. Layer 4 is divided into four layers, known as the granular layers. These layers receive feedforward input, and the magnocellular and parvocellular pathways remain segregated within these layers. Finally, deep layer 5 contains mainly pyramidal neurons which receive feedback from other areas. Layer 5 mainly projects to Layer 6 which has recurrent axons with layer 4 and the LGN. For a more in-depth summary of the cortical layer connections, see (Geyer et al., 2011; Goebel et al., 2012; Self et al., 2019; Thomson & Bannister, 2003). Overall, layers have distinct anatomical and physical properties (such as differences in energy consumption) (Goense et al., 2012) which coincide with functional differences, such as relaying feedforward and/or feedback signals. It is however, important to note that the division of layers is not clear-cut, with cell components spanning across layers (Larkum et al., 2018). Nevertheless, cortical models should encompass influences of laminar structure and function (Larkum et al., 2018) to account for this important feature of the cortex.

4.3 Traditional Methodologies

Classical investigation of the laminar cortex involves invasive techniques performed in-vivo in non-human primates, other mammals or rodents. These studies mainly rely on electrophysiological (Gilbert, 1977; Hubel & Wiesel, 1962) or optical-based methodologies which characterise the cortical response using neuronal firing rates or response to wavelengths of light, respectively. While their application has proven invaluable in aiding our understanding of cortical structure and function, these techniques have limitations. Animals do not have the same cognitive capabilities as humans and are not able to report perceptual processes, limiting the tasks that can be deployed in their experimentation. Additionally, these methodologies can have poor spatial resolution, particularly at the laminar level (De Martino et al., 2018) and have an extremely limited field-of-view, with an inability to measure the whole brain at once (Ugurbil, 2016). The nature of such investigations and the limit to non-human species means only potential inferences can be made about human cortical function based on their findings.

Human laminar profiling has traditionally been done either using slices of cortical tissue post-mortem or less frequently in-vivo in surgical patients. This brings about the same limitations as sedated or anaesthetised animal studies in that inferences can only be made about the structure of the cortex and not the functional capabilities. Initial advances in fMRI resolution allowed for early studies to demonstrate laminar profiling in humans (Cheng et al., 2001; Ress et al., 2007).

The goal is to bridge the gap between invasive techniques and the investigation of cortical function in humans (Kemper et al., 2018), allowing the awake and functioning human brain to be explored at such fine detail. Ultra-high field fMRI moves us a significant step closer to bridging that gap by enabling detailed investigation of neural processing in humans.

4.4 The need for non-invasive laminar profiling

While the invasive techniques typically used in animal studies and post-mortem investigations in humans have taught us a substantial amount about the layered cortex, inferences that can be drawn about human cognition are limited without the ability to study the laminar function of the cortex in-vivo. Being able to capture laminar activity non-invasively in humans is a key step in advancing our knowledge of the cortical processing underlying human cognition and behaviour.

Support is emerging within the literature for theories recognising the different input/output functions of neurons. For example, apical amplification (Phillips, 2017) proposes that the activation of a neuron depends on two sources of input; a bottom-up sensory input and a top-down contextual input which modulates the response to the feedforward information. This notion is supported by mechanisms such as BAC-propagation studied in-vivo in rodents (Larkum, 2013), whereby the firing rate of the cell is increased substantially by sufficient contextual information arriving from top-down cortical areas and targeting the cell’s dendrites. Much of the focus of neural processing is on the output of the cell body exclusively, without much consideration (if any) to the dendritic properties or influence. Since neurons span several cortical layers with their axons and dendrites, and findings are increasingly highlighting the importance of both types of input/output of the cell, the functional properties of each layer needs to be considered to gain an understanding of the functional property of the cortex itself.

Advances in MRI technology has led to the development of ultra-high field MRI. Most MRI scanners used within the neuroscientific community have a field strength of 3 Tesla (3T) which is considered high-field compared to standard hospital-grade scanners operating at 1.5T. However, there is a growing population of ultra-high field (7T and above) scanners being deployed within research facilities. Standard resolution (3-5mm isotropic resolution), high resolution (1-2mm isotropic) and ultra-high resolution (<1mm isotropic) can be used to investigate the neural response between different cortical areas, the response within a given cortical area, and the properties of neurons within that area, respectively, meaning different spatial scales are available with different resolutions (Olman & Yacoub, 2011). The advent of ultra-high field (UHF) has thus allowed the cortical function to be studied within a further dimension (cortical depth), which could aid in tracking the direction of flow within connections (De Martino et al., 2018). According to (Larkum et al., 2018), we still require an understanding of the input/output functions of most neurons and their role within the layered cortex. UHF MRI can provide microscopic scale insight, adding to our understanding of the meso- and macro-scopic function of the brain.

Increased resolution of UHF MRI allows for non-invasive laminar and columnar investigation of the cortex (De Martino et al., 2018), but how is this higher resolution achieved with an increase in field strength? At the physical level, increased field strength results in a linear increased polarisation of the proton signal, which forms the fundamental MRI signal (Olman & Yacoub, 2011), meaning a higher field strength leads to an increase in available MR signal. Functionally speaking, BOLD contrast also increases with field strength. Fundamentally, the underlying advantages of higher field strength lie in increased signal to noise ratio (SNR), increased spatial specificity and increased contrast to noise ratio (CNR) which ultimately leads to greater spatial resolution (De Martino et al., 2013; Olman & Yacoub, 2011). These enhanced parameters that result from increased magnetic field mean that UHF MRI can be used to acquire higher resolution functional and structural images.

The advantages that UHF MRI could potentially bring to the field of neuroscience mainly lie in using the increased resolution to explore the computational processing of the cortex at the columnar and laminar level. Theories highlighting a pivotal role of dendrite signalling can be tested non-invasively in humans using UHF MRI, with the ability to detect activity in dendrites that isn’t necessarily coupled with an action potential (Larkum et al., 2018). Furthermore, using higher resolution imaging could allow us to separately study the contributions of feedforward and feedback signals as well as excitatory/inhibitory inputs which may otherwise be coupled within a single voxel at a lower resolution (Goense et al., 2012). This could lead to crucial steps towards our understanding of the computational processing of the cortex. UHF eventually lends itself to novel opportunities such as investigation of processes that are uniquely human (De Martino et al., 2018) and the ability to compare the sub-millimetre components of many brains, giving insight into the variability of laminar structure and function.

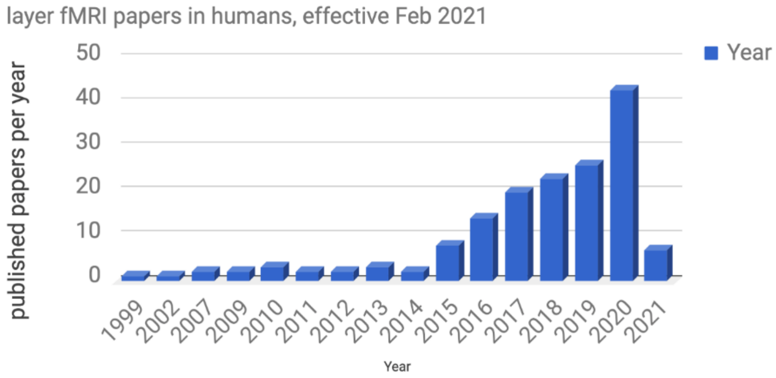


Figure 4.2 Layer fMRI publications per year. The number of layer-fMRI papers published per year has been steadily increasing since its advent. Figure obtained with permission from a layer-fMRI blog (https://layerfmri.com/).

4.5 Limitations of UHF MRI

4.5.1 General

While ultra-high field neuroimaging provides advantages in terms of the potential to investigate the cortex at a much more detailed spatial scale, the increased field strength doesn't come without some crucial limitations which must be considered. Olman & Yacoub (Olman & Yacoub, 2011) provide a review of the limitations surrounding UHF MRI and ultimately propose that higher resolutions are not always better. Pitfalls of MRI in general are often exacerbated in ultra-high resolution but it also highlights new challenges and considerations when running a UHF MRI experiment. The transition within the field to higher resolution seems slow and these limiting factors and challenges of UHF could play a role in this delay (De Martino et al., 2018).

Limitations can lie with acquiring data from human subjects, such as head movement within the scanner. When imaging with such fine spatial resolution, the range of motion tolerable within the data set is extremely limited, with head movement of a few millimetres able to cause considerable issues with data alignment and motion correction. Human subjects are also prone to fatigue or adaptation/practice effects. This limits the amount of data that can be collected within a scanning session. Increasing the number of samples collected can increase the SNR at UHF, however, this is caveated by the limitations of human capabilities. Specific absorption rate (SAR) is also an issue within MRI in general but the higher energy RF pulses at higher fields mean SAR limitations are more profound. UHF also results in increased distortion rates than lower field strengths and also leaves the images more prone to artifacts. See (Goense et al., 2016) for a more in-depth account of limitations encountered at UHF. These constraints of imaging at higher resolution must be considered within the context of the experimental aim in order to adequately account for their influences without compromising the potential of the experiment.

4.5.2 Neurovascular considerations

One of the major criticisms of MRI in general is its indirect link to neural processing. Functional MRI typically measures Blood Oxygen Level Dependent (BOLD) contrast, which reflects the metabolic demands of cortical processing. Other methodologies measure neural activation directly and thus this dissociation between the signal measured in MRI and the underlying neural activity is often highlighted as a key concern in its application and interpretation.

Neural activity leads to a local increase in oxygen consumption due to the metabolic demands of the neurons involved. This is enabled by an increase in both blood flow and blood volume (via vasodilation) to active regions of the cortex. The changes in blood flow and oxygen consumption rate leads to changes in levels of deoxyhaemoglobin, which is what forms the basis of the BOLD signal (Drew, 2019) and constitutes the haemodynamic response. Functional MRI signals are consequently tightly coupled with and dependent upon vascular properties. It is therefore crucial to gain a thorough understanding of the relationship between neural processing and vasculature dynamics in order to accurately determine what the BOLD response tells us about neuronal computations. See (Turner, 2016) for a review.

At standard (or high) resolution, voxel sizes above 1mm mean multiple features with different vascular properties can be confounded within one voxel. A single voxel above 1mm isotropic resolution could contain different blood vessels such as veins, arteries or capillaries or could capture signal from several layers or columns, which are known to differ in their haemodynamic profiles. These contributions can potentially be separated with ultra-high resolution, (Goense et al., 2012) therefore an understanding of their individual neurovascular dynamics is needed to interpret the signals achieved at higher resolution.

Cortical vasculature plays a more salient role in the generation and interpretation of BOLD signalling at higher resolution and is often considered a significant limitation to employing such a high spatial resolution. Being able to accurately interpret the haemodynamic response depends on our understanding of the neural and vascular components of the signal (Drew, 2019). One key area of concern lies in the large draining veins on the cortical surface. Blood enters the cortex via pial and intracortical arteries before reaching capillaries that have direct contact with neuronal tissue. It then drains through a series of veins eventually arriving in the large pial veins on the cortical surface (Turner, 2002). This results in a stronger BOLD signal amplitude in the superficial layers which decreases with layer depth (Dumoulin et al., 2018; Lawrence et al., 2019; Polimeni et al., 2010) as some BOLD signal is carried from deeper to superficial layers in this draining blood. Therefore activity in lower cortical layers will influence the signal in higher layers, (Markuerkiaga et al., 2016) meaning BOLD activation in superficial layers could partly reflect this influence of draining veins. Supply to the deeper layers is also therefore dependent on activation (and thus blood supply) to other layers (Goense et al., 2016). This draining effect also means that measured BOLD signal can reflect local changes in deoxyhaemoglobin as well as distant changes from draining veins (Turner, 2002).

The vertical nature of the vascular architecture also adds a temporal confound to the interpretation of the haemodynamic response. With temporal delays to some layers due to the vascular structure of the cortex, BOLD response onset therefore operates as a function of depth (Siero et al., 2011). The relationship between neural activity and haemodynamic response in sensory areas is considered to be explained by a linear convolution model, however this relationship is dependent upon factors such as the behavioural state or the cortical region (Drew, 2019). The temporal modulation of vasodilation of blood vessels can also lead to a blurring of the haemodynamic response with respect to the neuronal activity driving it (O'Herron et al., 2016). This particular temporal confound is beyond the temporal capabilities of MRI, however, being able to model the control of capillary dilation would enable local haemodynamic signals to be detected (Drew, 2019), which could have advantageous implications in terms of signal specificity.

The vascular architecture is not homogenous across layers. The mid-layers have a more concentrated network of blood vessels (Olman et al., 2012). Findings showing increased BOLD response in the mid-layers could simply be revealing a consequence of this expansion of vasculature rather than a true increase in signal due to neural activation (Koopmans et al., 2010). Synaptic transmission requires considerable energy, therefore the higher BOLD signal in these layers could also reflect the costly process of synaptic signalling (Logothetis, 2008).

The haemodynamic signal is also influenced by other factors relating to the vasculature of the cortex such as blood pooling effects that differ between layers (Heinzle et al., 2016). Variation in signal also exists due to the proximity of different types of vessels and their size and orientation (Ogawa et al., 1993), as well as the location of the water protons in terms of being intravascular or extravascular. These complex considerations regarding the vascular architecture and function of the cortex leas to drawbacks in being able to easily map the MRI signal to the underlying neural responses. I have only briefly touched on the multitude of vascular influences on the haemodynamic response and the complexities of the mechanisms behind them. For a more detailed account, see a review by (Uludağ & Blinder, 2018).

Despite these limitations, research has found distinct laminar differences in cortical processing, regardless of neurovascular confounds. Influences of the neurovascular system can be circumvented at the acquisition or analysis stage using a variety of techniques, discussed below. With adequate measures taken to limit these influences and research into understanding of the neurovascular taking place, it is possible to ensure the signals found reflect neuronal activity as opposed to biases induced by the cortical vasculature. Harnessing these haemodynamic profiles in the right way could also be beneficial for localisation of the response signal, giving greater insight into the localisation of the neural activity.

Pushing the spatial resolution to under 1 cubic millimetre may seem like the ideal standard to pursue, however, even voxels of this size could potentially contain input of thousands of neurons (Olman & Yacoub, 2011), making the disambiguation of underlying neural and vascular processes difficult. The importance of understanding the relationship between vascular and neuronal processes is not strictly limited to higher resolution and is a factor that is not yet resolved in lower resolution either (Logothetis, 2008), but higher resolution at least provides the opportunity to disambiguate these processes, even if we have not fully reached that point yet. Ultimately the trade-off between potential benefits of higher resolution and the complications that accompany it depends on question you want to ask. In a lot of instances, lower resolution may be sufficient, however, the ability to explore the cortex on mesoscopic scale undoubtedly benefits many areas of cognitive neuroscience.

4.6 UHF fMRI Acquisition & Analysis

Since its development, the application of UHF MRI has grown in popularity, with the number of studies using this technique rapidly increasing year by year. As such, high field brain mapping is a fast-developing field in human cognitive neuroscience (Larkum et al., 2018). This is driven by a steady influx of interest and resources along with a growing skill set within the neuroscientific community. Its expanding application has not been met without challenges and the advantages it brings to the field are not met without difficulties in the technicalities of data acquisition and analysis. In the same thread, the aforementioned limitations it encompasses can be overcome with appropriate acquisition and analysis parameters. The intricacies of each of these areas is inherently complex so I will only outline some of the key acquisition and analysis features which are currently deployed. For a more detailed discussion, see a review by (van der Zwaag et al., 2016).

4.6.1 Acquisition

To an extent, the limitations incurred at increased field strength can be minimised by accounting for such factors at the acquisition stage. By eliminating potential confounds a-priori, less needs to be done to deconvolve these factors from the resulting image signal. Several important parameters can be manipulated when acquiring UHF data which are not as influential at lower field strengths and therefore should be considered more carefully when running a higher resolution study.

One important aspect of UHF data acquisition is the imaging sequence that is used. Sequence design has the potential to improve both spatial and temporal resolution as well as increase the spatial specificity of the BOLD response (van der Zwaag et al., 2016). Gradient-Echo (GE) and Spin-Echo (SE) sequences are those most commonly used in ultra-high resolution fMRI, each encompassing specific benefits and drawbacks. Generally, GE sequences are more sensitive to blood oxygenation changes but are less spatially specific whereas SE methods are less sensitive than GE overall (Markuerkiaga et al., 2016) but less influenced by the vasculature of the cortex, reflecting the factors affecting T2 and T2* acquisition methods, respectively. GE-weighted sequences are more sensitive to the effects of draining veins and therefore can bias BOLD response signal to the pial surface. These draining effects lower the spatial specificity by contaminating BOLD response signal potentially far from its underlying neuronal source. Spin-Echo weighted measures are not immune to the effects of the vasculature architecture, with signal captured using SE sequencing affected by nearby capillaries (Goense et al., 2016). SE weighted imaging is also inefficient in terms of SNR, have less statistical power and can be problematic in their analysis in terms of motion correction and alignment with anatomical data (Koopmans & Yacoub, 2019) but SE based techniques have higher spatial specificity and are more laminar specific than GE. See Figure 4.3 for an illustration of a comparison between the two main acquisition types.

Overall, most UHR studies use T2* GE based sequencing due to the reduction in BOLD sensitivity incurred with T2-weighted sequences (van der Zwaag et al., 2016), but others have successfully used SE based sequences at the laminar level (Norris, 2012; Yacoub et al., 2007; Zhao et al., 2006). A combination of gradientand spin- echo imaging (3D GRASE) can also be used to counteract the drawbacks associated with each, but this too comes with limitations. It is important to note that the same acquisition method can still show differences in response profiles depending on other parameters such as field strength or TE (Goense et al., 2016) and different sequences can also show similar response profiles. Muckli and colleagues (Muckli et al., 2015) tested the same paradigm using T2 and T2* sequences and found no discernible differences between the two in terms of the output response.

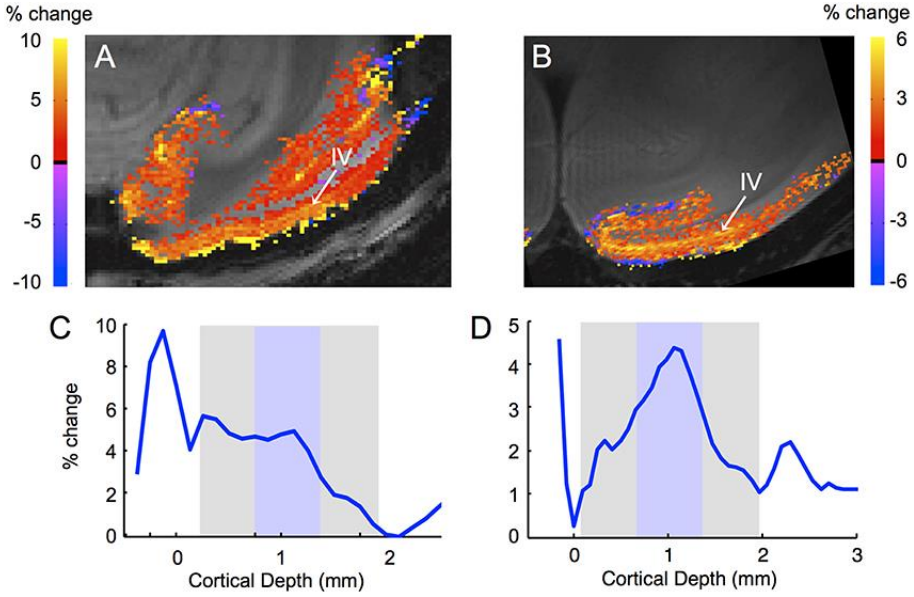


Figure 4.3 Comparison of Gradient-Echo (GE) and Spin-Echo (SE) sequences. GE (A) and SE (B) high-resolution sequences result in different laminar activations (C and D). GE BOLD response results in stronger activation at the cortical surface whereas SE BOLD signal is less sensitive to the veins on the cortical surface. Figure obtained with permission from (Goense et al., 2016) **using a Creative Commons Attribution (CC BY) licence.**

The resulting signal therefore depends on other contributing factors beyond the sequence employed. Koopmans and colleagues (Koopmans et al., 2011) have shown the influence of echo time (TE) on the resulting signal. They found sequences with short TEs resulted in the highest activation on the pial surface. Longer TEs however showed the most signal change in grey matter.

Slice coverage can also limit the capabilities of UHF imaging with temporal sampling rates needing to be taken into account to ensure sufficient time for the BOLD contrast to emerge. EPI and multiband methods can help achieve this as well as avoiding whole brain acquisition. Voxel size should also be considered as noise distortions can scale with voxel size, making smaller voxels more preferable (van der Zwaag et al., 2016).

Outside of neurovascular issues, other physiological factors need to be taken into account at the acquisition stage when imaging human subjects. Peripheral nerve stimulation and the levels of SAR (both of which differ in higher resolution imaging) need to be appropriately measured for human application. With human subjects being awake unlike animal studies, motion is an issue for MRI in general but is exacerbated at higher resolution. Sequence parameters are also sensitive to human behaviour inside the scanner, with different acquisition times influencing subject attention and motivation and ultimately, data quality.

So far, I have only mentioned acquisition within methodologies centred around the BOLD response. Increased SNR at higher fields allows for alternative imaging methods to be used (De Martino et al., 2018). Strategies harnessing other aspects of the neurovascular response are proving useful in negating some of the limitations of BOLD-focused techniques. Cerebral blood volume (CBV) and cerebral blood flow (CBF) methods can be used alone or in conjunction with BOLD profiling to capture different consequences of the neural response. Application of such techniques has proven fruitful (Huber et al., 2015, 2017; Poplawsky et al., 2019) and models have been developed to account for these haemodynamic factors. These techniques are sensitive to different physiological elements and can therefore provide different insight into cortical activity, not afforded by BOLD response. In fact, BOLD response has been considered inefficient for laminar investigation (Larkum et al., 2018) and so measurement of different vascular considerations could be more beneficial for layer-specific investigation.

4.6.2 Analysis

While the potential benefits of UHF MRI can be maximised through manipulation of acquisition parameters, the analysis of UHF data is also crucial in order to reap the benefits that higher resolution imaging can allow. Higher-order analysis is required of such data to account for any biases induced by vascular architecture, minimise confounds caused by participant behaviour and to attempt to circumvent any drawbacks that arise from different acquisition parameters.

Interpretation of raw signal changes is more challenging in higher resolution UHF imaging techniques. Therefore it is common practice to employ more advanced analysis techniques that go beyond the absolute signal change and instead consider relative signal change resulting from computational modelling (Huber et al., 2017). An example of this is performing classification-based analysis to determine the decodability of the signal within cortical layers (Muckli et al., 2015). Computational analysis also allows for physiological confounds to be modelled and ultimately removed from the relevant task-induced signal changes. These can be applied on BOLD-fMRI data or in conjunction with vascular profiling.

Removal of the superficial layers within data analysis can also be useful in consolidating the task signal and eliminating bias induced by pial veins or other vasculature effects (e.g., Polimeni, Fischl, Greve, & Wald, 2010) and improving spatial localisation. Vascular confounds have also been removed through the development of numerous different models (Moerel et al., 2018; Ugurbil, 2016). Other approaches focus on preserving the laminar profiles such as modelling based on equi-volume strategies (Kemper et al., 2018) or through temporal noise analysis (Koopmans et al., 2011). Analysis of UHR data has some flexibility by pooling along different dimensions of the cortical space allowing prioritisation of activity within different aspects of the cortical tissue to be examined (Kemper et al., 2018), such as laminar or regional-based analysis.

Different acquisition methods also pose different challenges at the analysis stage which can lead to different outcomes (Moerel et al., 2018). For example, T2-weighted imaging required a small FOV which can cause issues in motion-correction and registration between functional and anatomical data. Conversely, T2*-weighted imaging motion correction and co-registration is easier in higher-resolution data but is more susceptible to vascular-related issues (Koopmans & Yacoub, 2019).

Ultimately the acquisition and analysis strategies employed while running a high-resolution experiment can greatly affect the outcome of the data (Kashyap et al., 2018). There is no gold standard of practice which can eliminate all potential confounds and there will always be a trade-off between different advantages that coincide with different practices. Similarly there is no current standard practice in general, meaning there is no unanimous agreement on acquisition and analysis protocols as well as the ultimate interpretation of the data (Koopmans & Yacoub, 2019). UHF imaging is still in its relative infancy and therefore key developments are yet to be made to achieve such a consensus. It is therefore important to tailor the methodologies to suit the study aims and requirements in order to maximise the benefits that higher resolution can provide.

4.7 UHF in visual cortex

Understanding visual cognition is a key area of research within the neuroscientific community in general. The advent of ultra-high field fMRI has only strengthened this effort by providing the opportunity to investigate the visual cortex in a spatial scale not previously possible in human research. The higher resolution afforded by higher field MR imaging allows the neural computations and mechanisms operating in the visual cortex, highlighted by lower-resolution efforts, to be tested at a finer spatial scale. The higher resolution of UHF MRI also has the potential to tease apart processes of visual cognition operating within cortical columns or on different cortical layers. Furthermore, higher level visual processes have the opportunity to be explored in greater detail than previously afforded by lower-resolution and interactions that operate within the layers and columns across the visual hierarchy can be unveiled through UHF.

With the theoretical drive to seek to understand the mechanisms of visual processing on several spatial scales to gain insight into the function of the visual cortex (from the individual neuronal computations through the laminar and columnar operations, eventually to functional connectivity within the visual cortex itself and with other cortical areas), the visual cortex has already become one of the most explored cortical areas at UHF. From a practical perspective, the position of the visual cortex permits the use of local send and receive coils which maximise SNR and minimise implications of SAR (van der Zwaag et al., 2016), making it a good candidate for exploration at this resolution.

The application of UHF MRI has already proved fruitful in many aspects of visual perception and cognition research. Early UHF studies have successfully mapped ocular dominance columns (Cheng et al., 2001; Yacoub et al., 2007) and later orientation columns (Yacoub et al., 2008) within human V1. These findings are extended by the more recent work of (Dumoulin et al., 2018) who also find evidence of columnar organisation in V2 and V3. Outlining columnar function in human visual cortex is an important advancement of MRI research as previously cortical columns have only been exhibited architecturally in humans or functionally in animals using invasive techniques not suitable for human application. Columnar structure thus far has not been identifiable through standard (3T) resolution fMRI. Being able to investigate sensory areas at the columnar level could have important implications for the organisational and functional processing within these sub-units of cortical structure. The advantages of high resolution are however limited as such that the spatial resolution can focus on cortical depth in a small cortical region or be directed horizontally to capture surface-level activity over larger areas. Thus, functional connectivity between columns is a potential expansion of UHF imaging when the necessary developments are made.

Ultra-high field MRI has also been applied to the mapping of population receptive fields (PRFs). Higher resolution allows a more precise investigation of PRF properties in visual cortex, compared to the larger voxel sizes at lower resolutions (Zaretskaya, 2021) and also facilitates mapping across layers (Fracasso et al., 2016). Fracasso and colleagues found smaller PRF sizes in mid-layer cortex, relative to deep and superficial layers, revealing a U-shaped function of PRF size and cortical depth. Thus, there appears to be laminar distinctions of visuospatial mapping properties which suggests laminar-level distribution of sensory-driven input. Extending the high-field investigation of PRF properties to look at the influence of high-level processes, Klein and colleagues (Klein et al., 2018) found spatial attention to have a modulatory effect on PRF attraction. They revealed voluntary spatial attention attracts PRFs toward the attended location particularly in deep layers. This implies a top-down attentional affect that has a layer-specific profile.

In addition to revealing detailed attributes of various cortical maps and columns, high resolution fMRI has implications within higher level visual processing. Layer-specific attentional effects have been observed (Lawrence et al., 2019; Scheering et al., 2016) with different attention selectively modulating the activity in different cortical layers. Illusion studies at ultra-high resolution have also revealed layer-wise activation patterns employing classical illusory stimulation such as apparent motion (Schneider et al., 2019), Kanizsa shapes (Kok, Bains, et al., 2016), and illusions and mental imagery (Keogh et al., 2020); each of which point towards a layer dependent activation profile of visual awareness. Higher order processes outside of visual cortex such as working memory have also been shown to operate in a layer specific manner in primates (Koyano et al., 2016; Van Kerkoerle et al., 2017). Further investigation into the laminar profile of working memory is a therefore good candidate for ultra-high field fMRI exploration. In fact, the application of ultra-high field fMRI extends far beyond visual processing in general and can be applied to other sensory and non-sensory modalities as well as multisensory integration and functional connectivity. The progress and potential of which, however, is outside the scope of this thesis.

Layer-specific top-down modulation found in PRFs and other visual processes alludes to a segregation of top-down and bottom-up input within the layers of the visual cortex. There is a wealth of literature outlining the role of feedback pathways across the sensory cortex in general and in visual cortex in particular (Muckli & Petro, 2013; Petro et al., 2014). As such, feedback plays a key role in some of the central theories outlining cortical function (Phillips, 2017; Rao & Ballard, 1999), especially its integration with bottom-up input. The exact nature of this relationship between both types of input is still however, under debate and there is still much to learn about how and when these two processing streams converge within the visual hierarchy. Laminar fMRI could be a key tool in exploring the concurrent processing in visual cortex, specifically by mapping the layer-dependent segregation of feedback and feedforward signals.

Differences in the layer-wise activity between feedforward and feedback processing has been found across several studies (see Lawrence et al., 2019) and it is generally accepted that feedforward sensory input arrives in the mid-layers and feedback is most prominent in the superficial and deep layers. Two key studies have considered the implication of feedback within the layered cortex and have revealed seemingly conflicting findings.

In 2015, Muckli and colleagues (Muckli et al., 2015) partially occluded natural scene images to isolate feedback signals to a region of visual cortex in a high-resolution follow up to their lower-resolution finding that feedback signals can be decoded in unstimulated regions of V1 (Smith & Muckli, 2010). When the scene was presented in full, they found BOLD response signal to peak in the middle layers, where feedforward information is considered to project. This was revealed through MVPA and therefore scene information was most easily decoded in the mid-layers. However, when the image was partially occluded (that is, the bottom-right quadrant of the image was replaced with a uniform blank image, removing any scene information from the corresponding cortical region), activation peaked in the superficial layers. As such, information about the surrounding scene image was able to be decoded in these superficial layers, despite a lack of direct feedforward stimulation in this region. This implies feedback signals containing information about the surround are fed to the superficial layers of occluded V1 and reinforce the notion of a laminar distinction between the processing streams.

Kok and colleagues (Kok, Bains, et al., 2016) have also investigated the role of feedback in early visual cortex using a paradigm that also partially removes feedforward input. They utilised the Kanizsa illusion, in which contour configurations within the stimulus give rise to illusory shape edges and boundaries. As such, the illusion of a completed shape is thought to arise from expectations carried in feedback signals as these edges are perceived despite no direct bottom-up stimulation. Kok and colleagues find increased activity in the deep layers in regions of V1 corresponding to the illusory shape contours. Thus, implying that the feedback signals are projected to the deeper layers of visual cortex.

The findings of these two studies therefore differentially implicate both superficial and deep layers of cortex within feedback processing. These differences could simply reflect differences in acquisition and analysis techniques, with Muckli and colleagues focusing on the multivariate representation of the BOLD signal and Kok's study focusing on the univariate response. They also employed different sequencing to obtain the fMRI data, with Muckli opting for GE based EPI which is known to fall victim to the effects of large draining veins, driving a larger response in the superficial layers. Beyond technical disparities, the differences could also reflect distinctions in underlying processing resulting from the tasks themselves. A recent study by Bergmann and colleagues (Bergmann, 2019) found different activation profiles between different types of feedback signals; namely, mental imagery and illusory imagery. The feedback mechanisms used by Muckli and Kok could also have differing layer activation as they represent different aspects of visual processing (amodal and modal completion, respectively). Figure 4.4 illustrates the differences observed between the two studies.

The discrepancies noted here reveal how perhaps an important advantage of higher resolution investigation is that it may uncover key distinctions between findings that are currently clustered together on the basis of having compatible findings. Layer specific testing may expose differences currently concealed by lower-resolution limitations, leading to a deeper understanding of processes involved in visual perception.

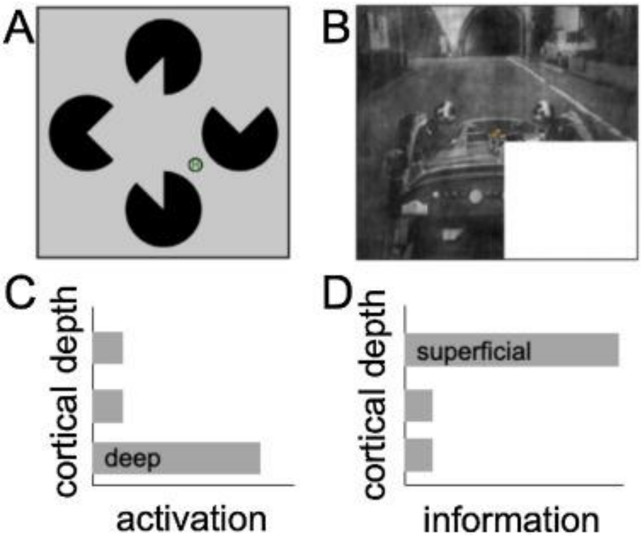


Figure 4.4 Deep and superficial layer effects. Comparison of (Kok, Bains, et al., 2016) **and** (Muckli et al., 2015) **findings obtained with permission from** (Zaretskaya, 2021) **using a Creative Commons Attribution (CC BY) licence. Stimuli used by Kok and colleagues is depicted by A and that by Muckli and colleagues depicted in B. Figures C and D demonstrate the depth dependent effects observed by each, respectively.**

4.8 Project proposal – Laminar investigation of top-down influence on the processing of degraded visual input

With the promising notion of delineating contributions of feedforward and feedback processing in visual cortex using layer-specific fMRI, we propose an experiment using ultra-high (7T) resolution to explore these fundamental cognitive processes further in this sensory modality.

4.8.1 Aim

This project aims to expand the findings of the two current projects which formed the basis of my PhD, and this thesis (Chapters 2 and 3). These projects addressed the fundamental research question of how top-down and bottom-up inputs interact to process degraded visual input. This was explored at the behavioural level in Chapter 2 and at the neuronal level in Chapter 3 using 3T fMRI. Here, I propose a project which explores this research question at the laminar level using ultra-high resolution (7T) fMRI.

The previous projects employed a partial-occlusion paradigm introduced by Smith & Muckli (Smith & Muckli, 2010) used by the original authors to isolate feedback signals within a region of visual cortex. In our paradigm, however, the occluded region contained degraded (low contrast) feedforward information in order to investigate how the surrounding context influences the processing of this input. The contextual information provided by the surround region is either consistent or inconsistent with the low-contrast (target region) information and therefore provides either accurate or inaccurate predictions about the content of the target region.

4.8.2 Rationale

Through our psychophysical experimental series, we found a reliable consistency effect in which consistent information in the surround enhances the recognition of low-contrast information in the target region and inconsistent information hinders performance. This effect somewhat persists when response bias is accounted for. Exploring this effect within the visual cortex revealed no distinct differences between the two surround types but does highlight the presence of feedback signals in occluded visual cortex. The presence of these signals is therefore revealed using conventional resolution fMRI; however, the content of these signals is still somewhat unknown. With no definitive differences between consistent and inconsistent feedback signals found at 3T, it prompts the need to explore these effects at higher resolution in an attempt to tease apart any differences which may be concealed by the lower resolution of 3T image acquisition. Employing this paradigm at 7T will allow us to explore whether there are any laminar signatures which could reveal an effect of consistent versus inconsistent top-down information.

Although our paradigm will replicate that used for the 3T experiment, I will briefly outline it here. Further details of the stimuli etc. can be found within Chapter 3. Our stimuli will involve grayscale images of natural scenes. These will be comprised of two regions; a high-contrast surround region which will encompass the majority of the visual scene and a low-contrast target region which will occupy the bottom-right quadrant. The two regions will either match (i.e., be consistent) or not match (i.e., be inconsistent). Subjects will be required to indicate the type of scene presented in the target region (e.g., a beach or mountain).

4.8.3 Experimental Approach

Experimentally, subjects will attend one scanning session at the Imaging Centre of Excellence (the University of Glasgow's 7T research facility, located at the Queen Elizabeth University Hospital, Glasgow). We will aim to collect data from around 20 subjects to ensure enough power in our results and to replicate sample sizes from previous 7T projects within the lab (e.g. (Morgan et al., 2019). Informed consent will be obtained, and we will follow all required protocol for scanning subjects at 7T. Although the full sequencing parameters are yet to be tested using this paradigm, we plan to use a GE-based sequence to acquire functional data of a cortical volume centred around the occipital cortex. Voxel size will be 0.8mm isotropic to achieve sufficient resolution for cortical-depth analysis. Analysis will be performed on each subject's V1 ROI which will be identified using both retinotopic mapping and target mapping stimulation. We will perform both univariate and multivariate analysis to examine differences in the amplitude of BOLD signal change between layers as well as potential differences in the response pattern.

4.8.4 Expected Outcomes

Firstly, we expect to find a laminar distinction between feedforward and feedback signals. Our paradigm will contain trials in which only feedforward or feedback information is presented and therefore we can assess the individual laminar profiles for each. As is consistently found within the literature, we expect feedforward information to peak in the mid-layers. Our feedforward input, however, is weak and therefore this effect may not be easily observable within our data. In terms of feedback, we expect to find activation primarily in the superficial layers, replicating the findings of (Muckli et al., 2015). Although feedback has also been observed in deep layers, the paradigm we employ here is similar to that used by Muckli and colleagues and therefore we expect our findings to reflect this similarity.

Secondly, we hope to observe differences at the laminar level between consistent and inconsistent feedforward/feedback signals. At 3T we do not find an overall distinction however, 3T resolution does not permit layer-wise effects to be investigated. As such, potential distinctions between these two consistency conditions will be distorted at 3T. The fact that we provide evidence of feedback influence on the processing of the feedforward input at the behavioural level and that feedback information drives a lot of the BOLD response alludes to an effect being present, just as yet uncovered. Conceptually, it seems inefficient of the visual system to represent these feedback signals in early visual cortex without capitalising on their content at this level. Results could reveal input of consistent versus inconsistent top-down information to different cortical layers or distinctions in the amount of signal they result in at the same or different depths.

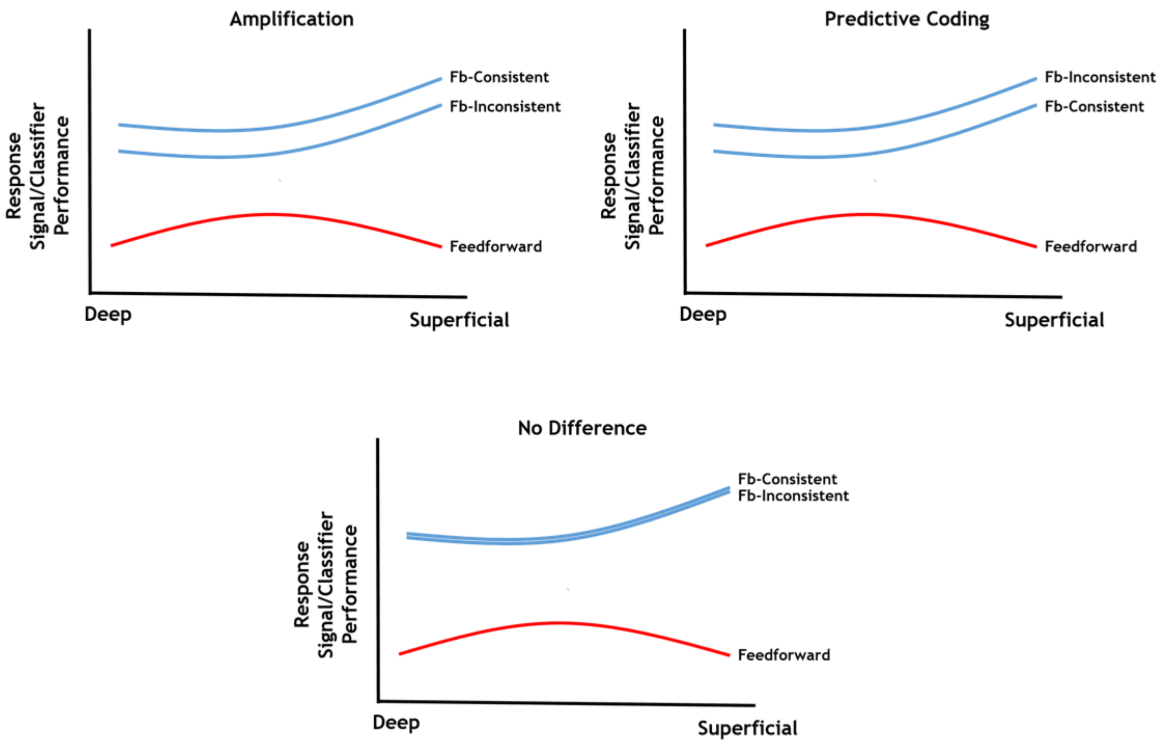


Figure 4.5 Schematic of expected results options. Potential results outcomes of the 7T experiment. We expect to find feedback signals to peak in the superficial layers and feedforward signals to peak in the mid layers. We also expect feedback signals to elicit a stronger response than the feedforward signals due to the degradation of the feedforward input. In the first scenario, we find an amplification of consistent information in which consistent feedback and feedforward signals result in a higher BOLD response signal (and/or higher classifier performance) than inconsistent fb/ff input. In the second scenario, inconsistent signals result in higher response signals (in line with predictive coding accounts of cortical function). Finally, in the third scenario we observe no differences between the consistent and inconsistent conditions.

4.8.5 Impact

It is hoped that exploring this paradigm at higher resolution will give insight into the cortical depth dependency of feedback and feedforward information. Not only will it allow contextual modulation to be investigated at a higher spatial resolution, but it will also add to the growing evidence of a laminar distinction between these two processing streams.

4.9 Conclusion

In sum, UHF fMRI allows the human brain to be explored on a new level which was previously only available using invasive animal studies. As such, it opens up the possibility to glean insight into human cortical function at a much finer spatial scale than previously achievable and may lead to findings about the computational capabilities of laminar and sub-laminar components.

Being able to image the cortex with such resolution is no doubt an invaluable advancement in the field, however we must now work to overcome its shortcomings to ensure to exploit the opportunities it lends. Furthermore, it is also important to build on the theoretical accounts of cortical function at this level to guide a-priori experimental design and hypotheses. Informed predictions about what to expect with UHR fMRI will require well-informed theories of cortical function (Larkum et al., 2018), in turn, knowledge gained from being able to explore the cortex on this level can be used to inform proposed theories and help better our understanding of the brain as a whole.

5 General Discussion

5.1 Aim

The work presented in this thesis set out to explore the influence of top-down feedback signals on the processing of feedforward information. More specifically, it looked at the role of contextual modulation within the processing of degraded visual input. It aimed to develop the findings of Smith & Muckli (Smith & Muckli, 2010) who found evidence of feedback signals within unstimulated early visual cortex. Through a series of experiments, we wanted to expand this paradigm to investigate whether these feedback signals can be used in the interpretation of visual input which is degraded and therefore difficult to perceive.

To achieve this, we developed what we termed the '*frosted occluder'* which gets its namesake from mimicking the perceptual effects of looking through a frosted screen. In such circumstances, visual information is perceptually available but is difficult to interpret. With the role of top-down signalling in contextual modulation sufficiently highlighted through several lines of research, we reasoned that providing contextual information that is easy to interpret will influence the ability to process the degraded or '*frosted*' input. In terms of direction of this effect, consistent contextual information has been shown to facilitate the processing of ambiguous information. Therefore, at the behavioural level we expected to find a facilitatory effect of consistency. The picture is a lot less clear at the neuronal level, with some theories suggesting inconsistency signals can be enhanced whereas others suggesting the opposite. Based on the concept of apical amplification, we wanted to test whether top-down predicted information can amplify dim, feedforward input.

5.2 Psychophysical Findings

We firstly ran a series of psychophysical experiments aiming to investigate different aspects of how top-down and bottom-up inputs integrate to influence the processing of weak, feedforward information. We utilised an adaptation of the partial occlusion paradigm introduced by Smith & Muckli in 2010 whereby a series of natural scenes are presented with the bottom right quadrant of the scene occluded. In our adaptation, the bottom right corner contained low-contrast (I.e., degraded) information that either matched or did not match the surrounding high-contrast scene, constituting the context. Subjects were asked to perform either a detection task (e.g., ‘*is there a scene present in the target region?’)* or a recognition task (e.g., ‘*did the target region contain a beach?’),* or in some instances, both tasks were performed in a within-subjects design.

In terms of detection, we considered detection rates between different consistency conditions and found that the consistency of the surrounding context (and therefore the reliability of the top-down predictions, by definition) did not influence subjects’ ability to detect whether or not a low-contrast scene was present. Detection rates were similar across consistency conditions but improved with increased visibility of the target region through an increase in its contrast level, reflecting a generic contrast effect rather than any effect of consistency.

For the recognition task, we looked at raw accuracy results as a measure of performance and found a significant advantage of consistent contextual information in the ability to determine the content of the low-contrast region. This effect was dichotomous, with inconsistent information hampering this ability. Performance was around chance level when no contextual information was present, reinforcing the influential ability of context.

We also analysed the data using Signal Detection Theory to account for any response bias within our consistency effect. Signal Detection was used to determine whether a consistency effect persists between consistency conditions while accounting for bias in responding in a particular direction. We still found evidence of increased sensitivity when the low-contrast target region was surrounded by contextually coherent information. Sensitivity was lower when the surrounding information was incoherent, indicating an influence of the predictability of top-down information on the interpretation of degraded information, independent of a tendency to respond in either direction.

In a second round of signal-detection-based analysis, we factored in a bias for subjects to respond according to the information presented in the surround. We reasoned that due to the ambiguous nature of the information in the target region and the difficulty in perceiving the low-contrast signal, subjects may depend on the information in the surround to make their judgement. This would result in a bias toward the surround, which could be interpreted as a perceptual bias or an influence of contextual modulation. The way in which the response components (hits and false alarms) are categorised within this analysis do not permit the division of responses into the consistency conditions. Therefore, we considered an overall indication of response bias and sensitivity within this framework. We did find a slight indication that subjects were inclined to presume the surround and target regions were consistent, particularly at the lower contrast levels where the target region is the most degraded. However, we did not find any evidence of extreme bias within any of our experiments and regardless of the tendency to respond according to the surround in some instances, we still find that subject sensitivity in being able to recognise the low-contrast scene was above chance-level.

We also manipulated properties of the contextual and feedforward information. We limited spatial frequency information in both the target and the surround in turn. When spatial frequencies were limited in the target region in addition to degradation of the contrast level, we found that low spatial frequencies were better at enabling subjects to identify the scene content, an effect which overrode the consistency effect we previously observed. When spatial frequencies were limited in the surround, contrary to previous findings implicating LSF information in contextual modulation, we found that when the context contained high spatial frequency information, subjects were better able to determine the low-contrast target information.

Overall, findings of our psychophysical experiment series highlight a congruency effect in contextual modulation which transpires as a behavioural response pattern. The tendency of subjects to occasionally respond according to the surround, that we observe through a slight response bias, may reflect a task-related mechanism employed when the feedforward input is too degraded. When the input is stronger, the surrounding context may then serve as a predictive mechanism employed by the top-down information stream. However, it may also suggest that the perceptual bias we accredit to contextual modulation is in part driven simply by a response bias in which subjects attend to the information in the surround and do not pay much attention to what is shown in the target region. We may therefore be attributing a behavioural outcome to the incorrect perceptual process, limiting the inferences that can be drawn about contextual modulation within this paradigm. Furthermore, we are unable to determine sensitivity within each condition independent of a surround bias (as we cannot test for an influence of consistency category when response bias for consistency between the target and surround regions are controlled for due to limitations in defining hit and false alarm rate required) which limits the interpretation of our findings somewhat. To circumvent this, SDT based analysis would need to be kept in mind when designing the experiment such that the design affords hit and false alarm rates to be determined within each condition. Alternatively, a design which reduces the ability to show bias towards the context would negate the need for this to be factored into the analysis (as in The Response Bias Experiment reported in 2.3.6.3).

Despite this, we provide evidence for the integration of top-down and bottom-up input when the bottom-up signal is degraded in the form of a behavioural pattern. Behaviour after all is how we interact with the environment and therefore we have demonstrated the outcome of contextual modulation in scene processing. It is also important, however, to understand the underlying neural processing of such an effect.

5.3 Neural Findings

We also explored our *frosted occluder* effect at the neuronal level. We wanted to test whether a consistency effect can be observed within the neural signal itself. That is, would we find differences between neural responses to feedback signals containing predictable (consistent) or unpredictable (inconsistent) information? More specifically, is top-down predictable information able to amplify dim feedforward input?

To achieve this, we applied the same *frosted occluder* paradigm to an fMRI experiment. We presented a series of visual scenes with the bottom right corner degraded. Instead of presenting the target region at a series of low contrast levels, we determined each subject’s perceptual threshold via a pre-screen detection task. We ran two fMRI experiments, with the target region being shown at either subjects’ 50% (i.e., lower-contrast) or 75% (i.e., higher-contrast) detection threshold. Therefore, the higher-contrast experiment was identical to the first except the target region and (therefore the feedforward information) was more visible. Again, the surround was shown in full-contrast and was either consistent or inconsistent with the scene in the target region. We also included a Target Only (in which no information was shown in the surround) and a Context Only (in which no information was shown in the target region) condition. Subjects completed a psychophysical recognition task within the scanner so that we could also acquire a behavioural response for each trial.

We analysed the neuronal responses within each consistency condition using both univariate and multivariate approaches. Our univariate analysis looked at the amplitude of signal change in response to the visual stimulation within the occluded region of visual cortex. Our multivariate analysis encompassed both cross-validation and cross-classification multivariate-pattern-analysis to reveal any differences in response patterns within occluded visual cortex.

In the lower-contrast experiment, we find a psychophysical advantage when the contextual information in the surround is consistent with the low-contrast target region, reinforcing the concept of a behavioural advantage of consistent context. In V1, the univariate analysis does not reveal any significant differences between predictable versus unpredictable information. Instead, we highlight the presence of feedback signals in occluded V1, with a significantly lower response signal when no contextual information is present. This is reflected at the multivariate level, with a cross-validation classification also highlighting the contribution of feedback signals to the response pattern; the classifier is only able to decode target-region information when feedback signals are present. Thus, it appears that the presence of feedback signals, rather than the content drives the response to degraded feedforward information in V1. V2 on the other hand, does appear to be influenced by the content of the feedback signals, with inconsistent information in the surround resulting in both a higher response amplitude and a higher classifier performance accuracy. This suggests that unpredictable information may drive a stronger response in occluded V2.

Motivated by the idea that the lack of differences in response to predictable versus unpredictable information in V1 could be due to feedforward information being too severely degraded that the feedback signals override any bottom-up influence, we increased the contrast of the target region in the higher-contrast experiment, therefore strengthening the feedforward signal. Increasing the feedforward information increased overall performance in the psychophysical task, however, rather than a consistency advantage, we observe more of an inconsistency disadvantage, with inconsistent information in the surround significantly reducing performance compared to either no or consistent information, which both resulted in similar performance accuracies.

In higher-contrast V1, we find a nearly significant higher response in the Inconsistent condition when only the correct trials are taken into account. When all trials are included, this distinction is non-significant at the univariate level, but we do find an advantage in the Inconsistent condition at the multivariate level within the cross-validation analysis. This inconsistency effect is also observed in a univariate analysis of V2 when only correct trials are considered.

Overall, increasing the strength of the feedforward input increases both psychophysical performance and the strength of the BOLD response amplitude. It also lessens the distinction between having feedback information present/absent, suggesting a greater contribution of feedforward information to the response signal when the input is stronger. Increased feedforward information also suggests an amplification of inconsistent information, modulated by task performance. This enhancement is also observed in the response pattern. Regardless, the key driver behind V1 response appears to be the presence of feedback signals, rather than their content. V2 on the other hand, appears to be more influenced by the consistency between the surround and target regions in both experiments, and demonstrates increased response to inconsistency. It would be interesting to model such contributions of these two early visual areas to the contextual modulation of visual processing. Differential weighting of features such as strength of feedback signal and content of feedback signal would need to be considered for each area to accurately determine the relative contribution of each factor within each area.

5.4 Layer-specific hypotheses

A natural next step in the course of this paradigm would be to explore these effects of contextual modulation at the laminar level. Recent advances in ultra-high-field fMRI allow high-resolution imaging of visual cortex which can be used to investigate cortical responses at a finer spatial scale than previously afforded, for example, within cortical columns or layers. This experiment lends itself to laminar investigation following a recent higher-resolution application of the Smith & Muckli partial occlusion paradigm (Muckli et al., 2015), who found laminar distinctions between feedforward and feedback information.

Applying this paradigm in which feedback signals are present in conjunction with degraded feedforward input may allow us to investigate how these two information processing streams are segregated and/or integrated within cortical layers. We expect to find a laminar difference between purely feedback and feedforward information but whether this distinction is maintained if both signals are combined, as in this paradigm is an interesting notion to investigate. Such an application will provide important insight into how bottom-up and top-down signals are processed, particularly in situations of perceptual uncertainty when top-down input is more crucially required.

5.5 Interpretation and future directions

Through this project, we demonstrate that context plays a role in both the behavioural and neural response to scene information. Through this contextual modulation, we highlight a salient role of top-down influences on the processing of degraded feedforward information. At the behavioural level, consistency between the context and available visual information is an important determinant of how this information is interpreted, as is the availability of such information. This is in line with the general literature which finds a dichotomous relationship between consistency/inconsistency and object/scene processing (Bar, 2004; Bar et al., 2006; Biederman et al., 1982).

Cortically, the picture is less clear, we find strong evidence to suggest that feedback signals are important for processing degraded feedforward information but how they are used and what information they convey is still to be determined. In our experiments, consistency of the feedback and feedforward inputs results in less of a distinction neuronally than behaviourally. We find some evidence of an enhancement of inconsistent information in the target region, suggesting a mechanism in line with predictive coding (Rao & Ballard, 1999) accounts of cortical processing as opposed to amplification of consistent signals as in apical amplification (Phillips, 2017), for example. We do also find some multivariate evidence of a segregation of feedback and feedforward signals in early visual cortex.

The psychophysical effect of consistency that we reliably find suggests that contextual modulation is a key component in the processing of degraded input and that this effect manifests as a behavioural response. Therefore, this effect must arise somewhere along the visual processing stream. The fact that our neuronal findings are less resounding does not mean that a consistency effect is not present at the neuronal level, the task is rather how we move forward and adapt our investigation in order to best expose such an effect. A further avenue for exploration could be to consider functional connectivity not only within the visual system (e.g. (Genç et al., 2016)) but across the whole brain within the context of this paradigm. Feedback signals arise from many regions in the brain (see Figure 1.2). Locating the source of these signals and exploring how the connections influence cortical function between areas will undoubtedly aid in the understanding of how feedback contributes to visual processing. Limiting our analysis to activity in primary visual cortex may prevent us from observing influences from other cortical regions that can help build a clearer picture of top-down and bottom-up integration.

Drawbacks of fMRI as a methodology could be limiting our search. The principles of fMRI lie within an indirect measure of cortical activity through its metabolic consequences. Perhaps a more direct measure of cortical activity will reveal top-down, and bottom-up interactions overlooked by fMRI limitations. Additionally, rather than constraints of fMRI in general, the lack of clear consistency effect could simply reflect our specific application. We focus our investigation in early visual cortex, particularly V1. The emergence of a consistency effect we observe in V2 could imply that an integration of signals occurs in higher visual areas. The role of V1 may be to relay the feedback and feedforward signals while their interaction may occur elsewhere, or perhaps contextual modulation is a result of functional connectivity which extends beyond visual cortex. Additionally, as mentioned, we may be limited by the resolution of our search and higher-resolution investigation may be the key in revealing top-down and bottom-up interactions within early processing stages.

One thing that is clear is that the notion of feedback influences on feedforward input processing is worth exploring in great detail in order to clarify how these processing streams allow us to visually perceive the world. The concept underlying this paradigm can be harnessed and adapted to build a clearer picture of top-down/bottom-up integration. For example, consistency can be manipulated in different ways beyond categorical features such as temporal aspects. We interact with scenes in a dynamic fashion and therefore temporal consistency is as crucial as spatial consistency in terms of scene processing. Adding a temporal element to the processing of degraded input could provide an interesting avenue to explore, as would other ways to degrade the visual input beyond lowering the contrast.

In terms of impact and how these findings contribute to the overall conceptual picture, we have provided evidence of a psychophysical effect of contextual modulation on scene categorisation, irrespective of a potential response bias. This supports previous notions of behavioural consistency effects in both objects and scenes and thus supports the idea that top-down predictions are used to facilitate the processing of bottom-up input, particularly in situations of ambiguity. There is an abundance of theories suggesting a distinction between the processing of predictable and unpredictable information within the visual hierarchy which is yet to be fully established at the cortical level , but we may need to redirect our focus to layer-wise comparisons in order to reveal where these distinctions exist.

The implications of these projects as they stand may have scope to extend beyond cortical processing of a healthy visual system. The role of feedback signals have been investigated within an experiment attempting to replicate central vision loss (Brown et al., 2021) to help understand the clinical implications of feedback signalling in visual cortex. Work has also looked at top-down signalling in those with vision loss (Masuda et al., 2010). Our findings may form a useful reference for the exploration of top-down modulatory effects when feedforward information is degraded, as is the case in clinical applications where vision has been compromised. Being able to build a picture of how the two processing streams function and integrate within healthy vision forms a foundation to understanding how feedback signalling may present when vision loss occurs. Being able to harness contextual information from within the visual system without the bottom-up input could have key implications for treating sight problems.

Nevertheless, despite the exciting prospect of real-world application of the findings beyond neuroscientific understanding of the cortex, the results from the PhD project as a whole contribute to the literature investigating top-down and bottom-up interactions within the visual system; a cortical function that is clearly salient in how we process our environment and one which is important to gain a full understanding of to human visual perception.

References

Abdelhack, M., & Kamitani, Y. (2018). Sharpening of Hierarchical Visual Feature Representations of Blurred Images. *Eneuro*. https://doi.org/10.1523/eneuro.0443-17.2018

Abrams, J., Nizam, A., & Carrasco, M. (2012). Isoeccentric locations are not equivalent: The extent of the vertical meridian asymmetry. *Vision Research*, *52*(1), 70–78. https://doi.org/10.1016/j.visres.2011.10.016

Afiki, Y., & Bar, M. (2020). Our need for associative coherence. *Humanities and Social Sciences Communications*, *7*(1), 1–11. https://doi.org/10.1057/s41599-020-00577-w

Alink, A., Schwiedrzik, C. M., Kohler, A., Singer, W., & Muckli, L. (2010). Stimulus Predictability Reduces Responses in Primary Visual Cortex. *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.3730-10.2010

Angelucci, A., & Bressloff, P. C. (2006). Chapter 5 Contribution of feedforward, lateral and feedback connections to the classical receptive field center and extra-classical receptive field surround of primate V1 neurons. *Progress in Brain Research*, *154*(SUPPL. A), 93–120. https://doi.org/10.1016/S0079-6123(06)54005-1

Angelucci, A., Levitt, J. B., Walton, E. J. S., Hupé, J. M., Bullier, J., & Lund, J. S. (2002). Circuits for local and global signal integration in primary visual cortex. *Journal of Neuroscience*, *22*(19), 8633–8646. https://doi.org/10.1523/jneurosci.22-19-08633.2002

Anzai, A., & DeAngelis, G. C. (2010). Neural computations underlying depth perception. *Current Opinion in Neurobiology*, *20*(3), 367–375. https://doi.org/10.1016/j.conb.2010.04.006

Bakin, J. S., Nakayama, K., & Gilbert, C. D. (2000). Visual responses in monkey areas V1 and V2 to three-dimensional surface configurations. *Journal of Neuroscience*, *20*(21), 8188–8198. https://doi.org/10.1523/jneurosci.20-21-

08188.2000

Ban, H., Yamamoto, H., Hanakawa, T., Urayama, S., Aso, T., Fukuyama, H., & Ejima, Y. (2013). Topographic Representation of an Occluded Object and the Effects of Spatiotemporal Context in Human Early Visual Areas. *The Journal of Neuroscience*. https://doi.org/10.1523/jneurosci.1455-12.2013

Bar, M. (2003). A cortical mechanism for triggering top-down facilitation in visual object recognition. *Journal of Cognitive Neuroscience*. https://doi.org/10.1162/089892903321662976

Bar, M. (2004). Visual objects in context. In *Nature Reviews Neuroscience*. https://doi.org/10.1038/nrn1476

Bar, M. (2007). The proactive brain: using analogies and associations to generate predictions. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2007.05.005

Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmidt, A. M., Dale, A. M., Hämäläinen, M. S., Marinkovic, K., Schacter, D. L., Rosen, B. R., & Halgren, E. (2006). Top-down facilitation of visual recognition. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.0507062103

Bar, M., & Ullman, S. (1996). Spatial context in recognition. *Perception*. https://doi.org/10.1068/p250343

Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., & Friston, K. J. (2012). Canonical Microcircuits for Predictive Coding. In *Neuron*. https://doi.org/10.1016/j.neuron.2012.10.038

Bastos, A. M., Vezoli, J., Bosman, C. A., Schoffelen, J. M., Oostenveld, R., Dowdall, J. R., DeWeerd, P., Kennedy, H., & Fries, P. (2015). Visual areas exert feedforward and feedback influences through distinct frequency channels. *Neuron*. https://doi.org/10.1016/j.neuron.2014.12.018

Benson, N. C., Butt, O. H., Datta, R., Radoeva, P. D., Brainard, D. H., & Aguirre, G. K. (2012). The retinotopic organization of striate cortex is well predicted by surface topology. *Current Biology*. https://doi.org/10.1016/j.cub.2012.09.014

Bergmann, J. (2019). *Short title: Two V1 feedback codes for internal experiences*. 1–38.

Berry, M. J., Brivanlou, I. H., Jordan, T. A., & Meister, M. (1999). Anticipation of moving stimuli by the retina. *Nature*, *398*(6725), 334–338. https://doi.org/10.1038/18678

Biderman, D., Shir, Y., & Mudrik, L. (2020). B or 13? Unconscious Top-Down Contextual Effects at the Categorical but Not the Lexical Level. *Psychological Science*, *31*(6), 663–677. https://doi.org/10.1177/0956797620915887

Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology*, *14*(2), 143–177. https://doi.org/10.1016/0010-0285(82)90007-X

Blasdel, G., & Lund, J. (1983). Termination of afferent axons in macaque striate cortex. *The Journal of Neuroscience*, *3*(7), 1389–1413.

Bonte, M., Parviainen, T., Hytönen, K., & Salmelin, R. (2006). Time course of top-down and bottom-up influences on syllable processing in the auditory cortex. *Cerebral Cortex*, *16*(1), 115–123. https://doi.org/10.1093/cercor/bhi091

Boycott, B., & Wassle, H. (1999). No TitleParallel Processing in the Mammalain Retina: the Proctor Lecture. *Investigative Ophthalmology and Visual Science*, *40*(1313–1327).

Brewer, A. A., Liu, J., Wade, A. R., & Wandell, B. A. (2005). Visual field maps and stimulus selectivity in human ventral occipital cortex. *Nature Neuroscience*, *8*(8), 1102–1109. https://doi.org/10.1038/nn1507

Brockmole, J. R., & Henderson, J. M. (2008). Prioritizing new objects for eye fixation in real-world scenes: Effects of object-scene consistency. *Visual Cognition*, *16*(2–3), 375–390. https://doi.org/10.1080/13506280701453623

Brown, H. D. H., Gouws, A. D., Vernon, R. J. W., Lawrence, S. J. D., Donnelly, G., Gill, L., Gale, R. P., Baseler, H. A., & Morland, A. B. (2021). Assessing functional reorganization in visual cortex with simulated retinal lesions. *Brain Structure and Function*, *226*(9), 2855–2867. https://doi.org/10.1007/s00429-021-02366-w

Brysbaert, M. (2019). How many participants do we have to include in properly powered experiments? A tutorial of power analysis with some simple guidelines. *Journal of Cognition*, *2*(1), 1–38.

Bubic, A., Yves von Cramon, D., & Schubotz, R. I. (2010). Prediction, cognition and the brain. *Frontiers in Human Neuroscience*, *4*(March), 1–15. https://doi.org/10.3389/fnhum.2010.00025

Budd, J. M. L. (1998). Extrastriate feedback to primary visual cortex in primates: A quantitative analysis of connectivity. *Proceedings of the Royal Society B: Biological Sciences*, *265*(1400), 1037–1044. https://doi.org/10.1098/rspb.1998.0396

Caddigan, E., Choo, H., Fei-Fei, L., & Beck, D. M. (2017). Categorization influences detection: A perceptual advantage for representative exemplars of natural scene categories. *Journal of Vision*. https://doi.org/10.1167/17.1.21

Carpenter, G. A., Grossberg, S., Systems, A., & Systems, N. (2016). Encyclopedia of Machine Learning and Data Mining. *Encyclopedia of Machine Learning and Data Mining*, 1–17. https://doi.org/10.1007/978-1-4899-7502-7

Chemla, S., Reynaud, A., Di Volo, M., Zerlaut, Y., Perrinet, L., Destexhe, A., & Chavane, F. (2019). Suppressive traveling waves shape representations of illusory motion in primary visual cortex of awake primate. *Journal of Neuroscience*, *39*(22), 4282–4298. https://doi.org/10.1523/JNEUROSCI.2792-

18.2019

Chen, G., Taylor, P. A., & Cox, R. W. (2017). Is the statistic value all we should care about in neuroimaging? *NeuroImage*, *15*(147), 952–959.

Chen, Y., Martinez-conde, S., Macknik, S. L., Bereshpolova, Y., Harvey, A., & Alonso, J. (2009). *NIH Public Access*. *11*(8), 974–982. https://doi.org/10.1038/nn.2147.TASK

Chen, Z., Denison, R. N., Whitney, D., & Maus, G. W. (2018). Illusory occlusion affects stereoscopic depth perception. *Scientific Reports*. https://doi.org/10.1038/s41598-018-23548-3

Chen, Z., Maus, G. W., Whitney, D., & Denison, R. N. (2017). Filling-in rivalry: Perceptual alternations in the absence of retinal image conflict. *Journal of Vision*. https://doi.org/10.1167/17.1.8

Cheng, K., Waggoner, R. A., & Tanaka, K. (2001). Cheng01. *Neuron*, *32*, 1–16. papers3://publication/uuid/67985648-9947-4952-AF2F-3B1D746D73C7

Cichy, R. M., Chen, Y., & Haynes, J. D. (2011). Encoding the identity and location of objects in human LOC. *NeuroImage*, *54*(3), 2297–2307. https://doi.org/10.1016/j.neuroimage.2010.09.044

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181–204. https://doi.org/10.1017/S0140525X12000477

Clavagnier, S., Falchier, A., & Kennedy, H. (2004). Long-distance feedback projections to area V1: Implications for multisensory integration, spatial awareness, and visual consciousness. *Cognitive, Affective and Behavioral Neuroscience*, *4*(2), 117–126. https://doi.org/10.3758/CABN.4.2.117

Connor, C. E., Egeth, H. E., & Yantis, S. (2004). Visual attention: Bottom-up versus top-down. *Current Biology*, *14*(19), 850–852. https://doi.org/10.1016/j.cub.2004.09.041

Corbett, J. E., & Carrasco, M. (2011). Visual performance fields: Frames of reference. *PLoS ONE*, *6*(9). https://doi.org/10.1371/journal.pone.0024470

Dagnelie, G. (2011). Visual prosthetics: Physiology, bioengineering, rehabilitation. In *Visual Prosthetics: Physiology, Bioengineering, Rehabilitation* (Issue January). https://doi.org/10.1007/978-1-4419-0754-7

Davenport, J. L., & Potter, M. C. (2004). Scene consistency in object and background perception. *Psychological Science*, *15*(8), 559–564. https://doi.org/10.1111/j.0956-7976.2004.00719.x

de Cesarei, A., & Loftus, G. R. (2011). Global and local vision in natural scene identification. *Psychonomic Bulletin and Review*, *18*(5), 840–847. https://doi.org/10.3758/s13423-011-0133-6

De Martino, F., Yacoub, E., Kemper, V., Moerel, M., Uludag, K., De Weerd, P., Ugurbil, K., Goebel, R., & Formisano, E. (2018). The impact of ultra-high field MRI on cognitive and computational neuroimaging. *NeuroImage*, *168*(October 2016), 366–382. https://doi.org/10.1016/j.neuroimage.2017.03.060

De Martino, F., Zimmermann, J., Muckli, L., Ugurbil, K., Yacoub, E., & Goebel, R. (2013). Cortical Depth Dependent Functional Responses in Humans at 7T: Improved Specificity with 3D GRASE. *PLoS ONE*, *8*(3), 30–32. https://doi.org/10.1371/journal.pone.0060514

Dekel, R., & Sagi, D. (2020). Interaction of contexts in context-dependent orientation estimation. *Vision Research*, *169*(March), 58–72. https://doi.org/10.1016/j.visres.2020.02.006

Denison, R. N., Piazza, E. A., & Silver, M. A. (2011). Predictive context influences perceptual selection during binocular rivalry. *Frontiers in Human Neuroscience*, *5*(NOVEMBER), 1–11. https://doi.org/10.3389/fnhum.2011.00166

DePauli, J. F. (1967). Book Review: Signal Detection Theory and Psychophysics .

Educational and Psychological Measurement, *27*(4), 921–922. https://doi.org/10.1177/001316446702700437

Dima, D. C., Perry, G., & Singh, K. D. (2018). Spatial frequency supports the emergence of categorical representations in visual cortex during natural scene perception. *NeuroImage*, *179*(June), 102–116. https://doi.org/10.1016/j.neuroimage.2018.06.033

DJ, F., & DC, V. E. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, *1*, 1–47. http://www.cogsci.ucsd.edu/~sereno/201/readings/04.03-MacaqueAreas.pdf

Dowling, J. . (1987). *The retina: an approachable part of the brain*. The Belknap Press of Harvard University Press.

Drew, P. J. (2019). Vascular and neural basis of the BOLD signal. *Current Opinion in Neurobiology*, *58*, 61–69. https://doi.org/10.1016/j.conb.2019.06.004

Dumoulin, S. O., Fracasso, A., van der Zwaag, W., Siero, J. C. W., & Petridou, N. (2018). Ultra-high field MRI: Advancing systems neuroscience towards mesoscopic human brain function. *NeuroImage*, *168*(January 2017), 345–357. https://doi.org/10.1016/j.neuroimage.2017.01.028

Duncan, R. O., & Boynton, G. M. (2003). Cortical magnification within human primary visual cortex correlates with acuity thresholds. *Neuron*, *38*(4), 659–671. https://doi.org/10.1016/S0896-6273(03)00265-4

Engel, A. ., Fries, P., & Singer, W. (2001). Dynamic predictions: Oscillations and synchrony in top-down processing. *Nature Reviews Neuroscience*, 2, 704–716.

Er, G., Pamir, Z., & Boyaci, H. (2020). Distinct patterns of surround modulation in V1 and hMT+. *NeuroImage*, *220*(July), 117084. https://doi.org/10.1016/j.neuroimage.2020.117084

Ernst, M. R., Triesch, J., & Burwick, T. (2019). Recurrent connectivity aids recognition of partly occluded objects. *ArXiv*, 1–9.

Faivre, N., Dubois, J., Schwartz, N., & Mudrik, L. (2019). Imaging object-scene relations processing in visible and invisible natural scenes. *Scientific Reports*, *9*(1), 1–13. https://doi.org/10.1038/s41598-019-38654-z

Fang, F., Kersten, D., & Murray, S. O. (2008). Perceptual grouping and inverse fMRI activity patterns in human visual cortex. *Journal of Vision*, *8*(7), 2–9. https://doi.org/10.1167/8.7.2

Fenske, M. J., Aminoff, E., Gronau, N., & Bar, M. (2006). Chapter 1 Top-down facilitation of visual object recognition: object-based and context-based contributions. *Progress in Brain Research*, *155 B*, 3–21. https://doi.org/10.1016/S0079-6123(06)55001-0

Ficco, L., Mancuso, L., Manuello, J., Teneggi, A., Liloia, D., Duca, S., Costa, T., Kovacs, G. Z., & Cauda, F. (2021). Disentangling predictive processing in the brain: a meta-analytic study in favour of a predictive network. *Scientific Reports*, *11*(1), 1–14. https://doi.org/10.1038/s41598-021-95603-5

Fracasso, A., Petridou, N., & Dumoulin, S. O. (2016). Systematic variation of population receptive field properties across cortical depth in human visual cortex. *NeuroImage*, *139*, 427–438. https://doi.org/10.1016/j.neuroimage.2016.06.048

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences*. https://doi.org/10.1098/rstb.2005.1622

Friston, K., Rigoli, F., Ognibene, D., Mathys, C., Fitzgerald, T., & Pezzulo, G. (2015). Active inference and epistemic value. *Cognitive Neuroscience*. https://doi.org/10.1080/17588928.2015.1020053

Genç, E., Schölvinck, M. L., Bergmann, J., Singer, W., & Kohler, A. (2016). Functional connectivity patterns of visual cortex reflect its anatomical organization. *Cerebral Cortex*, *26*(9), 3719–3731. https://doi.org/10.1093/cercor/bhv175

Gershman, S. J. (2019). What does the free energy principle tell us about the brain? *ArXiv*, 1–10.

Geyer, S., Weiss, M., Reimann, K., Lohmann, G., & Turner, R. (2011). Microstructural parcellation of the human cerebral cortex - from Bbrodmann's post-mortem map to in vivo mapping with high-field magnetic resonance imaging. *Frontiers in Human Neuroscience*, *5*(FEBRUARY), 1–7. https://doi.org/10.3389/fnhum.2011.00019

Gilbert, B. C. D. (1977). *Phypiol. (1977), 268,*. 391–421.

Gilbert, C. D., & Li, W. (2013). Top-down influences on visual processing. In *Nature Reviews Neuroscience*. https://doi.org/10.1038/nrn3476

Gilbert, C. D., & Sigman, M. (2007). Brain States: Top-Down Influences in Sensory Processing. In *Neuron*. https://doi.org/10.1016/j.neuron.2007.05.019

Goebel, R., Muckli, L., & Kim, D. S. (2012). Visual System. *The Human Nervous System*, 1301–1327. https://doi.org/10.1016/B978-0-12-374236-0.10037-9

Goense, J., Bohraus, Y., & Logothetis, N. K. (2016). fMRI at high spatial resolution implications for BOLD-models. *Frontiers in Computational Neuroscience*, *10*(Jun), 1–13. https://doi.org/10.3389/fncom.2016.00066

Goense, J., Merkle, H., & Logothetis, N. K. (2012). High-Resolution fMRI Reveals Laminar Differences in Neurovascular Coupling between Positive and Negative BOLD Responses. *Neuron*, *76*(3), 629–639. https://doi.org/10.1016/j.neuron.2012.09.019

Gollisch, T., & Meister, M. (2010). Eye Smarter than Scientists Believed: Neural Computations in Circuits of the Retina. *Neuron*, *65*(2), 150–164. https://doi.org/10.1016/j.neuron.2009.12.009

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *TINS*, *15*(1), 20–25.

Gosselin, F., & Schyns, P. G. (2003). Superstitious perceptions reveal properties of internal representations. *Psychological Science*, *14*(5), 505–509. https://doi.org/10.1111/1467-9280.03452

Gottesman, C. V., & Intraub, H. (1999). Wide-angle memories of close-up scenes: A demonstration of boundary extension. *Behavior Research Methods, Instruments, and Computers*, *31*(1), 86–93. https://doi.org/10.3758/BF03207697

Grabenhorst, F., Rolls, E. T., & Bilderbeck, A. (2008). How cognition modulates affective responses to taste and flavor: Top-down influences on the orbitofrontal and pregenual cingulate cortices. *Cerebral Cortex*, *18*(7), 1549–1559. https://doi.org/10.1093/cercor/bhm185

Green, C., & Hummel, J. E. (2006). Familiar interacting object pairs are perceptually grouped. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(5), 1107–1119. https://doi.org/10.1037/0096-1523.32.5.1107

Greenberg, A. S., Verstynen, T., Chiu, Y. C., Yantis, S., Schneider, W., & Behrmann, M. (2012). Visuotopic cortical connectivity underlying attention revealed with white-matter tractography. *Journal of Neuroscience*, *32*(8), 2773–2782. https://doi.org/10.1523/JNEUROSCI.5419-11.2012

Greene, M. R. (2013). Statistics of high-level scene context. *Frontiers in Psychology*, *4*(October), 1–31. https://doi.org/10.3389/fpsyg.2013.00777

Greve, D. N., & Fischl, B. (2009). Accurate and robust brain image alignment using boundary-based registration. *NeuroImage*, *48*(1), 63–72. https://doi.org/10.1016/j.neuroimage.2009.06.060

Grill-Spector, K., Kushnir, T., Hendler, T., Edelman, S., Itzchak, Y., & Malach, R. (1998). A sequence of object-processing stages revealed by fMRI in the human occipital lobe. *Human Brain Mapping*, *6*(4), 316–328. https://doi.org/10.1002/(SICI)1097-0193(1998)6:4<316::AID-HBM9>3.0.CO;2-6

Grill-Spector, K., & Malach, R. (2004a). The human visual cortex. *Annual Review of Neuroscience*, *27*, 649–677. https://doi.org/10.1146/annurev.neuro.27.070203.144220

Grill-Spector, K., & Malach, R. (2004b). THE HUMAN VISUAL CORTEX. *Annual Review of Neuroscience*. https://doi.org/10.1146/annurev.neuro.27.070203.144220

Groen, I. I. A., Jahfari, S., Seijdel, N., Ghebreab, S., Lamme, V. A. F., & Scholte, H. S. (2018). Scene complexity modulates degree of feedback activity during object detection in natural scenes. *PLoS Computational Biology*. https://doi.org/10.1371/journal.pcbi.1006690

Groen, I. I. A., Silson, E. H., & Baker, C. I. (2017). Contributions of low- and high-level properties to neural processing of visual scenes in the human brain. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1714). https://doi.org/10.1098/rstb.2016.0102

Grossberg, S. (2013). Adaptive Resonance Theory: How a brain learns to consciously attend, learn, and recognize a changing world. *Neural Networks*. https://doi.org/10.1016/j.neunet.2012.09.017

Hadjikhani, N. K., Dale, A. M., Liu, A. K., Cavanagh, P., & Tootell, R. B. H. (1998). FMRI of retinotopy and color sensitivity in human visual cortex. *NeuroImage*, *7*(4 PART II). https://doi.org/10.1016/s1053-8119(18)30875-9

Harris, J. J., Schwarzkopf, D. S., Song, C., Bahrami, B., & Rees, G. (2011). Contextual illusions reveal the limit of unconscious visual processing. *Psychological Science*, *22*(3), 399–405. https://doi.org/10.1177/0956797611399293

Hasson, U., Harel, M., Levy, I., & Malach, R. (2003). Large-scale mirrorsymmetry organization of human occipito-temporal object areas. *Neuron*, *37*(6), 1027–1041. https://doi.org/10.1016/S0896-6273(03)00144-2

Hawken, M. J., Parker, A. J., & Lund, J. S. (1988). Laminar organization and contrast sensitivity of direction-selective cells in the striate cortex of the Old World monkey. *Journal of Neuroscience*, *8*(10), 3541–3548. https://doi.org/10.1523/jneurosci.08-10-03541.1988

Haxby, J. V. (2012). Multivariate pattern analysis of fMRI: The early beginnings. *NeuroImage*, *62*(2), 852–855. https://doi.org/10.1016/j.neuroimage.2012.03.016

Haxby, J. V., Connolly, A. C., & Guntupalli, J. S. (2014). Decoding neural representational spaces using multivariate pattern analysis. *Annual Review of Neuroscience*, *37*, 435–456. https://doi.org/10.1146/annurev-neuro-062012-170325

Haynes, J. D., & Rees, G. (2005). Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nature Neuroscience*. https://doi.org/10.1038/nn1445

Heilbron, M., Richter, D., Ekman, M., Hagoort, P., & de Lange, F. P. (2020). Word contexts enhance the neural representation of individual letters in early visual cortex. *Nature Communications*, *11*(1), 1–11. https://doi.org/10.1038/s41467-019-13996-4

Heinzle, J., Koopmans, P. J., den Ouden, H. E. M., Raman, S., & Stephan, K. E. (2016). A hemodynamic model for layered BOLD signals. *NeuroImage*, *125*, 556–570. https://doi.org/10.1016/j.neuroimage.2015.10.025

Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. *Annual Review of Psychology*, *50*, 243–271. https://doi.org/10.1146/annurev.psych.50.1.243

Heywood, C. A., & Cowey, A. (1987). On the role of cortical area V4 in the discrimination of hue and pattern in macaque monkeys. *The Journal of*

Neuroscience : The Official Journal of the Society for Neuroscience, *7*(9), 2601–2617. https://doi.org/10.1523/jneurosci.07-09-02601.1987

Hinrichs, H., Scholz, M., Tempelmann, C., Woldorff, M. G., Dale, A. M., & Heinze, H. J. (2000). Deconvolution of event-related fMRI responses in fast-rate experimental designs: Tracking amplitude variations. *Journal of Cognitive Neuroscience*, *12*(SUPPL. 2), 76–89. https://doi.org/10.1162/089892900564082

Hock, H. S., Romanski, L., Galie, A., & Williams, C. S. (1978). Real-world schemata and scene recognition in adults and children. *Memory & Cognition*, *6*(4), 423–431. https://doi.org/10.3758/BF03197475

Hohwy, J., Roepstorff, A., & Friston, K. (2008). Predictive coding explains binocular rivalry: An epistemological review. *Cognition*. https://doi.org/10.1016/j.cognition.2008.05.010

Hollingworth, A., & Henderson, J. M. (1998). Does consistent scene context facilitate object perception? *Journal of Experimental Psychology: General*, *127*(4), 398–415.

Hsu, Y. F., Hämäläinen, J. A., & Waszak, F. (2014). Both attention and prediction are necessary for adaptive neuronal tuning in sensory processing. *Frontiers in Human Neuroscience*, *8*(MAR), 1–9. https://doi.org/10.3389/fnhum.2014.00152

Hubel, D. H., & Wiesel, T. (1962). AND FUNCTIONAL ARCHITECTURE IN THE CAT ' S VISUAL CORTEX From the Neurophysiolojy Laboratory , Department of Pharmacology central nervous system is the great diversity of its cell types and inter- receptive fields of a more complex type (Part I) and to. *Most*, 106–154.

Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *J. Physiol.*, *148*, 574–591.

Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology*, *195*(1), 215–243. https://doi.org/10.1113/jphysiol.1968.sp008455

Hubel, D. H., & Wiesel, T. N. (1974). Uniformity of monkey striate cortex: A parallel relationship between field size, scatter, and magnification factor. *Journal of Comparative Neurology*, *158*(3), 295–305. https://doi.org/10.1002/cne.901580305

Hubel, D. H., & Wiesel, T. N. (1998). Early exploration of the visual cortex. *Neuron*, *20*(3), 401–412. https://doi.org/10.1016/S0896-6273(00)80984-8

Hubel, D. H., Wiesel, T. N., & Stryker, M. P. (1978). Anatomical demonstration of orientation columns in macaque monkey. *Journal of Comparative Neurology*, *177*(3), 361–379. https://doi.org/10.1002/cne.901770302

Huber, L., Goense, J., Kennerley, A. J., Trampel, R., Guidi, M., Reimer, E., Ivanov, D., Neef, N., Gauthier, C. J., Turner, R., & Möller, H. E. (2015). Cortical lamina-dependent blood volume changes in human brain at 7T. *NeuroImage*, *107*, 23–33. https://doi.org/10.1016/j.neuroimage.2014.11.046

Huber, L., Handwerker, D. A., Jangraw, D. C., Chen, G., Hall, A., Stüber, C., Gonzalez-Castillo, J., Ivanov, D., Marrett, S., Guidi, M., Goense, J., Poser, B. A., & Bandettini, P. A. (2017). High-Resolution CBV-fMRI Allows Mapping of Laminar Activity and Connectivity of Cortical Input and Output in Human M1. *Neuron*, *96*(6), 1253-1263.e7. https://doi.org/10.1016/j.neuron.2017.11.005

Intraub, H., Gottesman, C. V, Willey, E. V, & Zuk, I. J. (1996). Boundary extension for briefly glimpsed photographs: Do common perceptual processes result in unexpected memory distortions? *Journal of Memory and Language*, *35*, 118–134.

Joo, S. J., Boynton, G. M., & Murray, S. O. (2012). Long-range, pattern-dependent contextual effects in early human visual cortex. *Current Biology*. https://doi.org/10.1016/j.cub.2012.02.067

K, O., & NG, K. (2000). Mental Imagery of Faces and Places Activates Corresponding Stimulus-Specific Brain Regions. *Journal of Cognitive Neuroscience*, *12*(6), 1013–1023.

Kashyap, S., Ivanov, D., Havlicek, M., Poser, B. A., & Uludağ, K. (2018). Impact of acquisition and analysis strategies on cortical depth-dependent fMRI. *NeuroImage*, *168*(May 2017), 332–344. https://doi.org/10.1016/j.neuroimage.2017.05.022

Katsuki, F., & Constantinidis, C. (2014). Bottom-up and top-down attention: Different processes and overlapping neural systems. *Neuroscientist*, *20*(5), 509–521. https://doi.org/10.1177/1073858413514136

Kauffmann, L., Ramanoël, S., & Peyrin, C. (2014). The neural bases of spatial frequency processing during scene perception. *Frontiers in Integrative Neuroscience*, *8*(MAY), 1–14. https://doi.org/10.3389/fnint.2014.00037

Kay, J., Floreano, D., & Phillips, W. A. (1998). Contextually guided unsupervised learning using local multivariate binary processors. *Neural Networks, 11*(1), 117–140. https://doi.org/10.1016/S0893-6080(97)00110-X

Kay, J. W., & Phillips, W. A. (2011). Coherent Infomax as a Computational Goal for Neural Systems. *Bulletin of Mathematical Biology*. https://doi.org/10.1007/s11538-010-9564-x

Kay, J. W., & Phillips, W. A. (2018). Contrasting information theoretic decompositions of modulatory and arithmetic interactions in neural information processing systems. *ArXiv*.

Keller, A. J., Roth, M. M., & Scanziani, M. (2020). Feedback generates a second receptive field in neurons of the visual cortex. *Nature*, *582*(7813), 545–549. https://doi.org/10.1038/s41586-020-2319-4

Keller, G. B., & Mrsic-Flogel, T. D. (2018). Predictive Processing: A Canonical Cortical Computation. *Neuron*, *100*(2), 424–435. https://doi.org/10.1016/j.neuron.2018.10.003

Kemper, V. G., De Martino, F., Emmerling, T. C., Yacoub, E., & Goebel, R. (2018). High resolution data analysis strategies for mesoscale human functional MRI at 7 and 9.4 T. *NeuroImage*, *164*(March 2017), 48–58. https://doi.org/10.1016/j.neuroimage.2017.03.058

Keogh, R., Bergmann, J., & Pearson, J. (2020). Cortical excitability controls the strength of mental imagery. *ELife*.

Kihara, K., & Takeda, Y. (2010). Time course of the integration of spatial frequency-based information in natural scenes. *Vision Research*, *50*(21), 2158–2162. https://doi.org/10.1016/j.visres.2010.08.012

Klein, B. P., Fracasso, A., van Dijk, J. A., Paffen, C. L. E., te Pas, S. F., & Dumoulin, S. O. (2018). Cortical depth dependent population receptive field attraction by spatial attention in human V1. *NeuroImage*, *176*(April), 301–312. https://doi.org/10.1016/j.neuroimage.2018.04.055

Kok, P., Bains, L. J., Van Mourik, T., Norris, D. G., & De Lange, F. P. (2016). Selective activation of the deep layers of the human primary visual cortex by top-down feedback. *Current Biology*. https://doi.org/10.1016/j.cub.2015.12.038

Kok, P., Brouwer, G. J., van Gerven, M. A. J., & de Lange, F. P. (2013). Prior Expectations Bias Sensory Representations in Visual Cortex. *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.0742-13.2013

Kok, P., Failing, M. F., & de Lange, F. P. (2014). Prior expectations evoke stimulus templates in the primary visual cortex. *Journal of Cognitive Neuroscience*. https://doi.org/10.1162/jocn_a_00562

Kok, P., Jehee, J. F. M., & de Lange, F. P. (2012). Less Is More: Expectation Sharpens Representations in the Primary Visual Cortex. *Neuron*. https://doi.org/10.1016/j.neuron.2012.04.034

Kok, P., Van Lieshout, L. L. F., & De Lange, F. P. (2016). Local expectation violations result in global activity gain in primary visual cortex. *Scientific*

Reports. https://doi.org/10.1038/srep37706

Koopmans, P. J., Barth, M., & Norris, D. G. (2010). Layer-specific BOLD activation in human V1. *Human Brain Mapping*, *31*(9), 1297–1304. https://doi.org/10.1002/hbm.20936

Koopmans, P. J., Barth, M., Orzada, S., & Norris, D. G. (2011). Multi-echo fMRI of the cortical laminae in humans at 7T. *NeuroImage*, *56*(3), 1276–1285. https://doi.org/10.1016/j.neuroimage.2011.02.042

Koopmans, P. J., & Yacoub, E. (2019). Strategies and prospects for cortical depth dependent T2 and T2* weighted BOLD fMRI studies. *NeuroImage*, *197*(February 2018), 668–676. https://doi.org/10.1016/j.neuroimage.2019.03.024

Koyano, K. W., Takeda, M., Matsui, T., Hirabayashi, T., Ohashi, Y., & Miyashita, Y. (2016). Laminar Module Cascade from Layer 5 to 6 Implementing Cue-to-Target Conversion for Object Memory Retrieval in the Primate Temporal Cortex. *Neuron*, *92*(2), 518–529. https://doi.org/10.1016/j.neuron.2016.09.024

Kravitz, D. J., Saleem, K. S., Baker, C. I., Ungerleider, L. G., & Mishkin, M. (2013). The ventral visual pathway: An expanded neural framework for the processing of object quality. *Trends in Cognitive Sciences*, *17*(1), 26–49. https://doi.org/10.1016/j.tics.2012.10.011

Kwisthout, J., Bekkering, H., & van Rooij, I. (2017). To be precise, the details don't matter: On predictive processing, precision, and level of detail of predictions. *Brain and Cognition*. https://doi.org/10.1016/j.bandc.2016.02.008

Larkum, M. (2013). A cellular mechanism for cortical associations: An organizing principle for the cerebral cortex. In *Trends in Neurosciences*. https://doi.org/10.1016/j.tins.2012.11.006

Larkum, M. E., Petro, L. S., Sachdev, R. N. S., & Muckli, L. (2018). A Perspective on Cortical Layering and Layer-Spanning Neuronal Elements. *Frontiers in Neuroanatomy*, *12*(July), 1–9. https://doi.org/10.3389/fnana.2018.00056

Lauer, T., Cornelissen, T. H. W., Draschkow, D., Willenbockel, V., & Võ, M. L. H. (2018). The role of scene summary statistics in object recognition. *Scientific Reports*. https://doi.org/10.1038/s41598-018-32991-1

Lawrence, S. J. D., Formisano, E., Muckli, L., & de Lange, F. P. (2019). Laminar fMRI: Applications for cognitive neuroscience. *NeuroImage*, *197*(May 2017), 785–791. https://doi.org/10.1016/j.neuroimage.2017.07.004

Lee, T. S., & Mumford, D. (2003). Hierarchical Bayesian inference in the visual cortex. *Journal of the Optical Society of America A*, *20*(7), 1434. https://doi.org/10.1364/josaa.20.001434

Lee, T. S., & Mumford, D. (2007). Hierarchical Bayesian inference in the visual cortex. *Journal of the Optical Society of America A*. https://doi.org/10.1364/josaa.20.001434

Levi, D. M. (2011). Visual crowding. *Current Biology*, *21*(18), R678–R679. https://doi.org/10.1016/j.cub.2011.07.025

Levitt, J. B., & Lund, J. S. (2002). The spatial extent over which neurons in macaque striate cortex pool visual signals. *Visual Neuroscience*, *19*(4), 439–452. https://doi.org/10.1017/S0952523802194065

Liang, H., Gong, X., Chen, M., Yan, Y., Li, W., & Gilbert, C. D. (2017). Interactions between feedback and lateral connections in the primary visual cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(32), 8637–8642. https://doi.org/10.1073/pnas.1706183114

Lindh, D., Sligte, I. G., Assecondi, S., Shapiro, K. L., & Charest, I. (2019). Conscious perception of natural images is constrained by category-related visual features. *Nature Communications*, *10*(1), 1–9. https://doi.org/10.1038/s41467-019-12135-3

Livingstone, M. S., & Hubel, D. H. (1987). Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. *Journal of Neuroscience*, *7*(11), 3416–3468. https://doi.org/10.1523/jneurosci.07-11-03416.1987

Logothetis, N. K. (2008). What we can do and what we cannot do with fMRI. *Nature*, *453*(7197), 869–878. https://doi.org/10.1038/nature06976

Lynn, S., Hoge, E., Fischer, L., Barrett, L. F., & Simon, N. (2014). Gender differences in oxytocin-associated disruption of decision bias during emotion perception. *Psychiatry Res.*, *219*(1), 198–203.

Lynn, S. K., & Barrett, L. F. (2014). “Utilizing” Signal Detection Theory. *Psychological Science*, *25*(9), 1663–1673. https://doi.org/10.1177/0956797614541991

Macaluso, E., & Doricchi, F. (2013). Attention and predictions: Control of spatial attention beyond the endogenous-exogenous dichotomy. *Frontiers in Human Neuroscience*, *7*(OCT), 75–80. https://doi.org/10.3389/fnhum.2013.00685

Mack, S. C., & Eckstein, M. P. (2011). Object co-occurrence serves as a contextual cue to guide and facilitate visual search in a natural viewing environment. *Journal of Vision*, *11*(9), 1–16. https://doi.org/10.1167/11.9.9

Maclean, K. A., Aichele, S. R., Bridwell, D. A., Mangun, G. R., Wojciulik, E., & Saron, C. D. (2009). Interactions between endogenous and exogenous attention during vigilance. *Attention, Perception, and Psychophysics*, *71*(5), 1042–1058. https://doi.org/10.3758/APP.71.5.1042

Macmillan, N. A., & Creelman, C. D. (2005). *Detetction Theory A User’s Guide*.

Major, G., Larkum, M. E., & Schiller, J. (2013). Active Properties of Neocortical Pyramidal Neuron Dendrites. *Annual Review of Neuroscience*. https://doi.org/10.1146/annurev-neuro-062111-150343

Manita, S., Suzuki, T., Homma, C., Matsumoto, T., Odagawa, M., Yamada, K., Ota, K., Matsubara, C., Inutsuka, A., Sato, M., Ohkura, M., Yamanaka, A., Yanagawa, Y., Nakai, J., Hayashi, Y., Larkum, M. E., & Murayama, M. (2015). A Top-Down Cortical Circuit for Accurate Sensory Perception. *Neuron*, *86*(5), 1304–1316. https://doi.org/10.1016/j.neuron.2015.05.006

Mannion, D. J., Kersten, D. J., & Olman, C. A. (2015). Scene coherence can affect the local response to natural images in human V1. *European Journal of Neuroscience*. https://doi.org/10.1111/ejn.13082

Markov, N. T., Vezoli, J., Chameau, P., Falchier, A., Quilodran, R., Huissoud, C., Lamy, C., Misery, P., Giroud, P., Ullman, S., Barone, P., Dehay, C., Knoblauch, K., & Kennedy, H. (2014). Anatomy of hierarchy: Feedforward and feedback pathways in macaque visual cortex. *Journal of Comparative Neurology*, *522*(1), 225–259. https://doi.org/10.1002/cne.23458

Markuerkiaga, I., Barth, M., & Norris, D. G. (2016). A cortical vascular model for examining the specificity of the laminar BOLD signal. *NeuroImage*, *132*, 491–498. https://doi.org/10.1016/j.neuroimage.2016.02.073

Marzecová, A., Widmann, A., SanMiguel, I., Kotz, S. A., & Schröger, E. (2017). Interrelation of attention and prediction in visual processing: Effects of task-relevance and stimulus probability. *Biological Psychology*, *125*(2017), 76–90. https://doi.org/10.1016/j.biopsycho.2017.02.009

Masuda, Y., Horiguchi, H., Dumoulin, S. O., Furuta, A., Miyauchi, S., Nakadomari, S., & Wandell, B. A. (2010). Task-dependent V1 responses in human retinitis pigmentosa. *Investigative Ophthalmology and Visual Science*, *51*(10), 5356–5364. https://doi.org/10.1167/iovs.09-4775

Maximilian Riesenhuber, & Tomaso Poggio. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience*, 2(11), 1019–1025.

Mayer, B. J., Biol, M. C., Stemmler, M., Usher, M., & Nieburt, E. (2018). *Lateral Interactions in Primary Visual Cortex : A Model Bridging Physiology and Psychophysics Author (s): Martin Stemmler , Marius Usher and Ernst Niebur Published by : American Association for the Advancement of Science*

Stable URL : http://www.jstor.or. *269*(5232), 1877–1880.

Mehrpour, V., Martinez-Trujillo, J. C., & Treue, S. (2020). Attention amplifies neural representations of changes in sensory input at the expense of perceptual accuracy. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-020-15989-0

Meijer, G. T., Montijn, J. S., Pennartz, C. M. A., & Lansink, C. S. (2017). Audiovisual Modulation in Mouse Primary Visual Cortex Depends on Cross-Modal Stimulus Configuration and Congruency. *The Journal of Neuroscience*. https://doi.org/10.1523/jneurosci.0468-17.2017

Miller, M. B., & Gazzaniga, M. S. (1998). Creating false memories for visual scenes. *Neuropsychologia*, *36*(6), 513–520. https://doi.org/10.1016/S0028-3932(97)00148-6

Millin, R., Arman, A. C., Chung, S. T. L., & Tjan, B. S. (2014). Visual crowding in V1. *Cerebral Cortex*, *24*(12), 3107–3115. https://doi.org/10.1093/cercor/bht159

Milner, A. D., & Goodale, M. A. (2008). Two visual systems re-viewed. *Neuropsychologia*, *46*(3), 774–785. https://doi.org/10.1016/j.neuropsychologia.2007.10.005

Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983). *Mishkin et al. (1983) Object Vision and Spacial vision*.

Moerel, M., De Martino, F., Kemper, V. G., Schmitter, S., Vu, A. T., Uğurbil, K., Formisano, E., & Yacoub, E. (2018). Sensitivity and specificity considerations for fMRI encoding, decoding, and mapping of auditory cortex at ultra-high field. *NeuroImage*, *164*(March 2017), 18–31. https://doi.org/10.1016/j.neuroimage.2017.03.063

Mohan, H., Verhoog, M. B., Doreswamy, K. K., Eyal, G., Aardse, R., Lodder, B. N., Goriounova, N. A., Asamoah, B., B. Brakspear, A. B. C., Groot, C., Van Der Sluis, S., Testa-Silva, G., Obermayer, J., Boudewijns, Z. S. R. M., Narayanan, R. T., Baayen, J. C., Segev, I., Mansvelder, H. D., & De Kock, C. P. J. (2015). Dendritic and axonal architecture of individual pyramidal neurons across layers of adult human neocortex. *Cerebral Cortex*, *25*(12), 4839–4853. https://doi.org/10.1093/cercor/bhv188

Morgan, A. T., Petro, L. S., & Muckli, L. (2019). Scene representations conveyed by cortical feedback to early visual cortex can be described by line drawings. *Journal of Neuroscience*, *39*(47), 9410–9423. https://doi.org/10.1523/JNEUROSCI.0852-19.2019

Motter, B. C. (1993). Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. *Journal of Neurophysiology*, *70*(3), 909–919. https://doi.org/10.1152/jn.1993.70.3.909

Mu, T., & Li, S. (2013). The neural signature of spatial frequency-based information integration in scene perception. *Experimental Brain Research*, *227*(3), 367–377. https://doi.org/10.1007/s00221-013-3517-1

Muckli, L., De Martino, F., Vizioli, L., Petro, L. S., Smith, F. W., Ugurbil, K., Goebel, R., & Yacoub, E. (2015). Contextual Feedback to Superficial Layers of V1. *Current Biology*. https://doi.org/10.1016/j.cub.2015.08.057

Muckli, L., Kohler, A., Kriegeskorte, N., & Singer, W. (2005). Primary visual cortex activity along the apparent-motion trace reflects illusory perception. *PLoS Biology*, *3*(8). https://doi.org/10.1371/journal.pbio.0030265

Muckli, L., & Petro, L. S. (2013). Network interactions: Non-geniculate input to V1. In *Current Opinion in Neurobiology*. https://doi.org/10.1016/j.conb.2013.01.020

Mumford, D. (1991). On the computational architecture of the neocortex - I. The role of the thalamo-cortical loop. *Biological Cybernetics*. https://doi.org/10.1007/BF00202389

Mumford, D., & Mumford, D. (1992). On the computational architecture of the neocortex. II. The role of cortico-cortical loops. *Biological Cybernetics*.

Murray, M. M., Foxe, D. M., Javitt, D. C., & Foxe, J. J. (2004). Setting boundaries: Brain dynamics of modal and amodal illusory shape completion in humans. *Journal of Neuroscience*, *24*(31), 6898–6903. https://doi.org/10.1523/JNEUROSCI.1996-04.2004

Murray, S. O., Boyaci, H., & Kersten, D. (2006). The representation of perceived angular size in human primary visual cortex. *Nature Neuroscience*. https://doi.org/10.1038/nn1641

Nanay, B. (2018). The Importance of Amodal Completion in Everyday Perception. *I-Perception*, *9*(4). https://doi.org/10.1177/2041669518788887

Norris, D. G. (2012). Spin-echo fMRI: The poor relation? *NeuroImage*, *62*(2), 1109–1115. https://doi.org/10.1016/j.neuroimage.2012.01.003

O'Herron, P., Chhatbar, P. Y., Levy, M., Shen, Z., Schramm, A. E., Lu, Z., & Kara, P. (2016). Neural correlates of single-vessel haemodynamic responses in vivo. *Nature*, *534*(7607), 378–382. https://doi.org/10.1038/nature17965

Ogawa, S., Menon, R. S., Tank, D. W., Kim, S. G., Merkle, H., Ellermann, J. M., & Ugurbil, K. (1993). Functional brain mapping by blood oxygenation level-dependent contrast magnetic resonance imaging. A comparison of signal characteristics with a biophysical model. *Biophysical Journal*, *64*(3), 803–812. https://doi.org/10.1016/S0006-3495(93)81441-3

Oliva, A., & Torralba, A. (2001). 2001 Gist Features. *Ijcv*, *42*(3), 145–175.

Oliva, A., & Torralba, A. (2003). Statistics of natural image categories. *Network: Computation in Neural Systems*, *14*(3), 391–412.

Oliva, A., & Torralba, A. (2006). Chapter 2 Building the gist of a scene: the role of global image features in recognition. *Progress in Brain Research*, *155 B*(March), 23–36. https://doi.org/10.1016/S0079-6123(06)55002-2

Oliva, A., & Torralba, A. (2007). The role of context in object recognition. In

Trends in Cognitive Sciences. https://doi.org/10.1016/j.tics.2007.09.009

Olman, C. A., Harel, N., Feinberg, D. A., He, S., Zhang, P., Ugurbil, K., & Yacoub, E. (2012). Layer-specific fmri reflects different neuronal computations at different depths in human V1. *PLoS ONE*, *7*(3). https://doi.org/10.1371/journal.pone.0032536

Olman, C. A., & Yacoub, E. (2011). High-Field fMRI for Human Applications: An Overview of Spatial Resolution and Signal Specificity. *The Open Neuroimaging Journal*, *5*(Suppl 1), 74–89. https://doi.org/10.2174/1874440001105010074

Palmer, tephen E. (1975). The effects of contextual scenes on the identification of objects. *Memory & Cognition*, *3*(5), 519–526. https://doi.org/10.3758/BF03197524

Palomero-Gallagher, N., & Zilles, K. (2019). Cortical layers: Cyto-, myelo-, receptor- and synaptic architecture in human cortical areas. *NeuroImage*, *197*(August 2017), 716–741. https://doi.org/10.1016/j.neuroimage.2017.08.035

Pan, H., Zhang, S., Pan, D., Ye, Z., Yu, H., Ding, J., Wang, Q., Sun, Q., & Hua, T. (2021). Characterization of Feedback Neurons in the High-Level Visual Cortical Areas That Project Directly to the Primary Visual Cortex in the Cat. *Frontiers in Neuroanatomy*, *14*(January), 1–14. https://doi.org/10.3389/fnana.2020.616465

Paneri, S., & Gregoriou, G. G. (2017). Top-down control of visual attention by the prefrontal cortex. Functional specialization and long-range interactions. *Frontiers in Neuroscience*, *11*(SEP), 1–16. https://doi.org/10.3389/fnins.2017.00545

Panichello, M. F., Cheung, O. S., & Bar, M. (2013). Predictive feedback and conscious visual experience. In *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2012.00620

Pelli, D. G., & Bex, P. (2013). Measuring contrast sensitivity. *Vision Research*. https://doi.org/10.1016/j.visres.2013.04.015

Petro, L. S., Paton, A. T., & Muckli, L. (2017). Contextual modulation of primary visual cortex by auditory signals. In *Philosophical Transactions of the Royal Society B: Biological Sciences*. https://doi.org/10.1098/rstb.2016.0104

Petro, L. S., Smith, F. W., Schyns, P. G., & Muckli, L. (2013). Decoding face categories in diagnostic subregions of primary visual cortex. *European Journal of Neuroscience*, *37*(7), 1130–1139. https://doi.org/10.1111/ejn.12129

Petro, L. S., Vizioli, L., & Muckli, L. (2014). Contributions of cortical feedback to sensory processing in primary visual cortex. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2014.01223

Phillips, W. A. (2017). Cognitive functions of intracellular mechanisms for contextual amplification. *Brain and Cognition*. https://doi.org/10.1016/j.bandc.2015.09.005

Phillips, W. A., Clark, A., & Silverstein, S. M. (2015). On the functions, mechanisms, and malfunctions of intracortical contextual modulation. *Neuroscience and Biobehavioral Reviews*, *52*, 1–20. https://doi.org/10.1016/j.neubiorev.2015.02.010

Phillips, W. A., Larkum, M. E., Harley, C. W., & Silverstein, S. M. (2016). The effects of arousal on apical amplification and conscious state. *Neuroscience of Consciousness*. https://doi.org/10.1093/nc/niw015

Põder, E. (2008). Crowding with detection and coarse discrimination of simple visual features. *Journal of Vision*, *8*(4), 1-6. https://doi.org/10.1167/8.4.24

Polimeni, J. R., Fischl, B., Greve, D. N., & Wald, L. L. (2010). Laminar analysis of 7T BOLD using an imposed spatial activation pattern in human V1. *NeuroImage*, *52*(4), 1334–1346. https://doi.org/10.1016/j.neuroimage.2010.05.005

Poplawsky, A. J., Fukuda, M., Kang, B. man, Kim, J. H., Suh, M., & Kim, S. G. (2019). Dominance of layer-specific microvessel dilation in contrast-enhanced high-resolution fMRI: Comparison between hemodynamic spread and vascular architecture with CLARITY. *NeuroImage*, *197*(January 2017), 657–667. https://doi.org/10.1016/j.neuroimage.2017.08.046

Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General*, *109*(2), 160–174. https://doi.org/10.1037/0096-3445.109.2.160

Pratte, M. S., & Tong, F. (2014). Spatial specificity of working memory representations in the early visual cortex. *Journal of Vision*, *14*(3), 1–12. https://doi.org/10.1167/14.3.22

Qiu, A., Rosenau, B. J., Greenberg, A. S., Hurdal, M. K., Barta, P., Yantis, S., & Miller, M. I. (2006). Estimating linear cortical magnification in human primary visual cortex via dynamic programming. *NeuroImage*, *31*(1), 125-138. https://doi.org/10.1016/j.neuroimage.2005.11.049

Raizada, R. D. S., & Grossberg, S. (2001). Context-sensitive binding by the laminar circuits of V1 and V2: A unified model of perceptual grouping, attention, and orientation contrast. *Visual Cognition*, *8*(3–5), 431–466. https://doi.org/10.1080/13506280143000070

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*. https://doi.org/10.1038/4580

Raslear, T. G. (1985). Perceptual bias and response bias in temporal bisection. *Perception & Psychophysics*, *38*(3), 261–268. https://doi.org/10.3758/BF03207153

Ress, D., Glover, G. H., Liu, J., & Wandell, B. (2007). Laminar profiles of functional activity in the human brain. *NeuroImage*, *34*(1), 74-84. https://doi.org/10.1016/j.neuroimage.2006.08.020

Revina, Y. (2021). *Enhances Contextual Feedback and Feedforward. February.*

Rideaux, R., & Welchman, A. E. (2019). Contextual effects on binocular matching are evident in primary visual cortex. *Vision Research*, *159*(April), 76–85. https://doi.org/10.1016/j.visres.2019.04.001

Rockland, K. S. (2017). Anatomy of the Cerebral Cortex. In *The Cerebral Cortex in Neurodegenerative and Neuropsychiatric Disorders: Experimental Approaches to Clinical Issues*. https://doi.org/10.1016/B978-0-12-801942-9.00001-X

Rockland, K. S., & Lund, J. S. (1983). Intrinsic laminar lattice connections in primate visual cortex. *Journal of Comparative Neurology*, *216*(3), 303–318. https://doi.org/10.1002/cne.902160307

Ronconi, L., Bertoni, S., & Bellacosa Marotti, R. (2016). The neural origins of visual crowding as revealed by event-related potentials and oscillatory dynamics. *Cortex*, *79*, 87–98. https://doi.org/10.1016/j.cortex.2016.03.005

Rossit, S., McAdam, T., Mclean, D. A., Goodale, M. A., & Culham, J. C. (2013). FMRI reveals a lower visual field preference for hand actions in human superior parieto-occipital cortex (SPOC) and precuneus. *Cortex*, *49*(9), 2525–2541. https://doi.org/10.1016/j.cortex.2012.12.014

Saarela, T. P., Sayim, B., Westheimer, G., & Herzog, M. H. (2009). Global stimulus configuration modulates crowding. *Journal of Vision*, *9*(2), 1–11. https://doi.org/10.1167/9.2.5

Sadaghiani, S., Uğurbil, K., & Uludağ, K. (2009). Neural activity-induced modulation of BOLD poststimulus undershoot independent of the positive signal. *Magnetic Resonance Imaging*, *27*(8), 1030–1038. https://doi.org/10.1016/j.mri.2009.04.003

Saito, T., & Sadoshima, J. (2016). Acting without seeing: Eye movements reveal visual processing without awareness Miriam Spering & Marisa Carrasco. *Trends Neurosci.*, *116*(8), 1477–1490.

https://doi.org/10.1016/j.tins.2015.02.002.Acting

Sasaki, H., Todorokihara, M., Ishida, T., Miyachi, J., Kitamura, T., & Aoki, R. (2006). Effect of noise on the contrast detection threshold in visual perception. *Neuroscience Letters*. https://doi.org/10.1016/j.neulet.2006.08.054

Sauvé, G., Harmand, M., Vanni, L., & Brodeur, M. B. (2017). The probability of object-scene co-occurrence influences object identification processes. *Experimental Brain Research*. https://doi.org/10.1007/s00221-017-4955-y

Scheering, R., Koopmans, P. J., Van Mourik, T., Jensen, O., & Norris, D. G. (2016). The relationship between oscillatory EEG activity and the laminar-specific BOLD signal. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(24), 6761–6766. https://doi.org/10.1073/pnas.1522577113

Schmidtmann, G., Logan, A. J., Kennedy, G. J., Gordon, G. E., & Loffler, G. (2015). Distinct lower visual field preference for object shape. *Journal of Vision*, *15*(5), 1–15. https://doi.org/10.1167/15.5.18

Schneider, M., Kemper, V. G., Emmerling, T. C., De Martino, F., & Goebel, R. (2019). Columnar clusters in the human motion complex reflect consciously perceived motion axis. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(11), 5096–5101. https://doi.org/10.1073/pnas.1814504116

Schyns, P. G., & Oliva, A. (1994). From blobs to boundary edges: Evidence for Time- and Spatial-Scale-Dependent Scene Recognition. *Psychological Science*, *5*(4), 195–200. https://doi.org/10.1111/j.1467-9280.1994.tb00500.x

Sekuler, R., Watamaniuk, N. ., & Blake, R. (2002). Perception of Visual Motion. In *Stevens' Handbook of Experimental PSychology, Sensation and Perception* (Third, pp. 121–177).

Self, M. W., van Kerkoerle, T., Goebel, R., & Roelfsema, P. R. (2019). Benchmarking laminar fMRI: Neuronal spiking and synaptic activity during top-down and bottom-up processing in the different layers of cortex. *NeuroImage*, *197*(March 2017), 806–817. https://doi.org/10.1016/j.neuroimage.2017.06.045

Selker, R., van den Bergh, D., Criss, A. H., & Wagenmakers, E. J. (2019). Parsimonious estimation of signal detection models from confidence ratings. *Behavior Research Methods*, *51*(5), 1953–1967. https://doi.org/10.3758/s13428-019-01231-3

Semedo, J. D., Jasper, A. I., Zandvakili, A., Krishna, A., Aschner, A., Machens, C. K., Kohn, A., & Yu, B. M. (2022). Feedforward and feedback interactions between visual cortical areas use different population activity patterns. *Nature Communications*, *13*(1), 1–14. https://doi.org/10.1038/s41467-022-28552-w

Serences, J. T., & Boynton, G. M. (2007). Feature-Based Attentional Modulations in the Absence of Direct Visual Stimulation. *Neuron*. https://doi.org/10.1016/j.neuron.2007.06.015

Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., Rosen, B. R., & Tootell, R. B. H. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science*, *268*(5212), 889–893. https://doi.org/10.1126/science.7754376

Shipp, S. (2016). Neural elements for predictive coding. In *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2016.01792

Siero, J. C. W., Petridou, N., Hoogduin, H., Luijten, P. R., & Ramsey, N. F. (2011). Cortical depth-dependent temporal dynamics of the BOLD response in the human brain. *Journal of Cerebral Blood Flow and Metabolism*, *31*(10), 1999–2008. https://doi.org/10.1038/jcbfm.2011.57

Singer, W. (2013). Cortical dynamics revisited. *Trends in Cognitive Sciences*, *17*(12), 616–626. https://doi.org/10.1016/j.tics.2013.09.006

Smith, F. W., & Muckli, L. (2010). Nonstimulated early visual areas carry information about surrounding context. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.1000233107

Smith, M. L., Gosselin, F., & Schyns, P. G. (2012). Measuring internal representations from behavioral and brain data. *Current Biology*, *22*(3), 191–196. https://doi.org/10.1016/j.cub.2011.11.061

Sobel, K. V, & Blake, R. (2002). How context influences predominance during binocular rivalry. *Perception*, *31*(7), 813–824.

Spratling, M. W. (2010). Predictive Coding as a Model of Response Properties in Cortical Area V1. *Journal of Neuroscience*. https://doi.org/10.1523/jneurosci.4911-09.2010

Straube, S., & Fahle, M. (2011). Visual detection and identification are not the same: Evidence from psychophysics and fMRI. *Brain and Cognition*, *75*(1), 29–38. https://doi.org/10.1016/j.bandc.2010.10.004

Sugita, Y. (1999). Grouping of image fragments in primary visual cortex. *Nature*. https://doi.org/10.1038/45785

Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends in Cognitive Sciences*, *13*(9), 403–409. https://doi.org/10.1016/j.tics.2009.06.003

Swets, J. A., Green, D. M., Getty, D. J., & Swets, J. B. (1978). Signal detection and identification at successive stages of observation. *Perception & Psychophysics*, *23*(4), 275–289. https://doi.org/10.3758/BF03199711

Teufel, C., & Nanay, B. (2017). How to (and how not to) think about top-down influences on visual perception. *Consciousness and Cognition*. https://doi.org/10.1016/j.concog.2016.05.008

Thomson, A., & Bannister, A. (2003). Interlaminar connections in the neocortex. *Cerebral Cortex*, *13*(1), 5–14.

Tobimatsu, S., Tomoda, H., & Kato, M. (1995). Parvocellular and magnocellular contributions to visual evoked potentials in humans: stimulation with chromatic and achromatic gratings and apparent motion. *Journal of the Neurological Sciences*, *134*(1–2), 73–82. https://doi.org/10.1016/0022-510X(95)00222-X

Torralba, A. (2003). Contextual priming for object detection. *International Journal of Computer Vision*, *53*(2), 169–191. https://doi.org/10.1023/A:1023052124951

Trampel, R., Bazin, P. L., Pine, K., & Weiskopf, N. (2019). In-vivo magnetic resonance imaging (MRI) of laminae in the human cortex. *NeuroImage*, *197*(April 2017), 707–715. https://doi.org/10.1016/j.neuroimage.2017.09.037

Turner, R. (2002). How much codex can a vein drain? Downstream dilution of activation-related cerebral blood oxygenation changes. *NeuroImage*, *16*(4), 1062–1067. https://doi.org/10.1006/nimg.2002.1082

Turner, R. (2016). Uses, misuses, new uses and fundamental limitations of magnetic resonance imaging in cognitive science. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1705). https://doi.org/10.1098/rstb.2015.0349

Ugurbil, K. (2016). What is feasible with imaging human brain function and connectivity using functional magnetic resonance imaging. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1705). https://doi.org/10.1098/rstb.2015.0361

Uludağ, K., & Blinder, P. (2018). Linking brain vascular physiology to hemodynamic response in ultra-high field MRI. *NeuroImage*, *168*(February 2017), 279–295. https://doi.org/10.1016/j.neuroimage.2017.02.063

van der Zwaag, W., Schäfer, A., Marques, J. P., Turner, R., & Trampel, R. (2016). Recent applications of UHF-MRI in the study of human brain function and structure: a review. *NMR in Biomedicine*, *29*(9), 1274–1288.

https://doi.org/10.1002/nbm.3275

van Kemenade, B. M., Wilbertz, G., Müller, A., & Sterzer, P. (2020). Non-stimulated regions in early visual cortex encode the contents of conscious visual perception. *BioRxiv*. https://doi.org/10.1101/2020.11.13.381269

Van Kerkoerle, T., Self, M. W., & Roelfsema, P. R. (2017). Layer-specificity in the effects of attention and working memory on activity in primary visual cortex. *Nature Communications*, *8*. https://doi.org/10.1038/ncomms13804

Vetter, P., Edwards, G., & Muckli, L. (2012). Transfer of predictive signals across saccades. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2012.00176

Vezoli, J., Magrou, L., Goebel, R., Wang, X. J., Knoblauch, K., Vinck, M., & Kennedy, H. (2021). Cortical hierarchy, dual counterstream architecture and the importance of top-down generative networks. *NeuroImage*, *225*(September 2020), 117479. https://doi.org/10.1016/j.neuroimage.2020.117479

Walther, D. B., Caddigan, E., Fei-Fei, L., & Beck, D. M. (2009). Natural scene categories revealed in distributed patterns of activity in the human brain. *Journal of Neuroscience*, *29*(34), 10573–10581. https://doi.org/10.1523/JNEUROSCI.0559-09.2009

Wandell, B. A., Glover, G. H., Eduardo-Jose, C., Shadlen, M. N., Engel, S. A., David E. Rumelhart, & Wandell. (1994). fMRI of human visual cortex. *Nature*, *369*(June), 525.

Wandell, B. A., & Winawer, J. (2011). Imaging retinotopic maps in the human brain. In *Vision Research*. https://doi.org/10.1016/j.visres.2010.08.004

Wang, L., Weng, X., & He, S. (2012). Perceptual grouping without awareness: Superiority of Kanizsa triangle in breaking interocular suppression. *PLoS ONE*, *7*(6), 1–6. https://doi.org/10.1371/journal.pone.0040106

Wässle, H. (2004). Parallel processing in the mammalian retina. *Nature Reviews Neuroscience*, *5*(10), 747–757. https://doi.org/10.1038/nrn1497

White, R., Engelen, G., & Uljee, I. (2016). Theory and Consequences. *Modeling Cities and Regions As Complex Systems*, *14*(1), 13–42. https://doi.org/10.7551/mitpress/9780262029568.003.0002

Wibral, M., Priesemann, V., Kay, J. W., Lizier, J. T., & Phillips, W. A. (2017). Partial information decomposition as a unified approach to the specification of neural goal functions. *Brain and Cognition*. https://doi.org/10.1016/j.bandc.2015.09.004

Williams, M. A., Baker, C. I., Op De Beeck, H. P., Mok Shim, W., Dang, S., Triantafyllou, C., & Kanwisher, N. (2008). Feedback of visual object information to foveal retinotopic cortex. *Nature Neuroscience*. https://doi.org/10.1038/nn.2218

Wörgötter, F., & Eysel, U. T. (2000). Context, state and the receptive fields of striatal cortex cells. In *Trends in Neurosciences*. https://doi.org/10.1016/S0166-2236(00)01632-5

Yacoub, E., Harel, N., & Uğurbil, K. (2008). High-field fMRI unveils orientation columns in humans. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(30), 10607–10612. https://doi.org/10.1073/pnas.0804110105

Yacoub, E., Shmuel, A., Logothetis, N., & Uğurbil, K. (2007). Robust detection of ocular dominance columns in humans using Hahn Spin Echo BOLD functional MRI at 7 Tesla. *NeuroImage*, *37*(4), 1161–1177. https://doi.org/10.1016/j.neuroimage.2007.05.020

Young, M. P. (2000). Architecture of visual cortex and inferential processes in vision. *Spatial Vision*, *13*(2–3), 137–146. https://doi.org/10.1163/156856800741162

Zagha, E., Casale, A. E., Sachdev, R. N. S., McGinley, M. J., & McCormick, D. A.

(2013). Motor cortex feedback influences sensory processing by modulating network state. *Neuron*. https://doi.org/10.1016/j.neuron.2013.06.008

Zaretskaya, N. (2021). Zooming-in on higher-level vision: High-resolution fMRI for understanding visual perception and awareness. *Progress in Neurobiology*, *November 2020*, 101998. https://doi.org/10.1016/j.pneurobio.2021.101998

Zeki, S. (2015). Area V5—a microcosm of the visual brain. *Frontiers in Integrative Neuroscience*, *9*(APRIL), 1–18. https://doi.org/10.3389/fnint.2015.00021

Zhao, F., Wang, P., Hendrich, K., Ugurbil, K., & Kim, S. G. (2006). Cortical layer-dependent BOLD and CBV responses measured by spin-echo and gradient-echo fMRI: Insights into hemodynamic regulation. *NeuroImage*, *30*(4), 1149–1160. https://doi.org/10.1016/j.neuroimage.2005.11.013

Zhaoping, L. (2014). Understanding Vision Theory, Models, and Data. *Oxford University Press*, *44*(8), 1689–1699.