

**Mapping invariance in the ventral occipital  
temporal pathway**

Chris Racey

Ph.D.

University of York  
Department of Psychology

2012

## Abstract

This thesis is concerned with the organisation of the human ventral occipital temporal pathway, part of the brain involved in the recognition of objects and processing of scenes. The architecture of this system has been extensively investigated with functional magnetic resonance imaging (fMRI) which reveals the engagement of parts of the brain in different tasks. Conventional analyses have been used to identify discrete, circumscribed brain regions which appear to be specialised for processing particular categories of stimuli, such as faces or landscapes. In characterising the function of these regions, an important factor in previous research has been to establish whether or not they are invariant to more basic sensory properties such as size or 3D orientation, which change as places and objects are encountered from different points of view. A more recent suggestion is that the ventral occipital temporal pathway, rather than being comprised of discrete regions, is better conceived of as a continuous topographical map, or multiple overlapping maps. Within this framework, fundamental computational requirements, such as the need to establish different forms of invariance, may provide overarching organising principles governing the spatial arrangement of processing within ventral visual cortex. This more nuanced view of the functional architecture of the visual system demands a more nuanced view of invariance. Through a series of experiments using fMRI in conjunction with a parametric adaptation paradigm, this thesis investigates the invariance of neural representations as a continuous variable that can be explicitly quantified rather than an all-or-none phenomenon. This novel approach reveals structure and symmetry normally hidden in conventional analyses. Invariance is found to vary by degree throughout the ventral surface of the brain depending on the nature of the stimulus and the types of processing required. Different forms of invariance can be dissociated from one another in terms of distinct spatial patterns of parametric adaptation that different forms of visual change generate.

# Table of contents

<b>Abstract</b> .....	<b>i</b>
<b>Table of contents</b> .....	<b>ii</b>
<b>List of figures</b> .....	<b>vi</b>
<b>List of tables</b> .....	<b>viii</b>
<b>Acknowledgements</b> .....	<b>ix</b>
<b>Declaration</b> .....	<b>x</b>
<b>1 The ventral occipital temporal pathway</b> .....	<b>1</b>
1.1 Parcellating the ventral stream.....	5
1.2 Organising principles of the VOT.....	7
1.3 Invariance as an organising principle.....	11
1.4 Using fMRI to investigate high level representations.....	12
1.4.1 MVPA.....	13
1.4.2 fMR-adaptation.....	14
1.5 Probing representational space.....	18
1.6 Summary and aims.....	21
<b>2 Processing of landscapes in place selective cortex: effects of parametric manipulation of viewpoint</b> .....	<b>23</b>
2.1 Introduction.....	23
2.1.1 Visual scene processing in the brain.....	26
2.1.2 The functional role of the PPA.....	27
2.1.3 The functional role of the RSC.....	29
2.1.4 Summary and aims.....	31
2.2 Methods.....	34
2.2.1 Participants.....	34
2.2.2 Imaging parameters.....	34
2.2.3 Landscape stimuli.....	34
2.2.4 Behavioural task.....	35
2.2.5 Localiser scan.....	35
2.2.6 Adaptation scan.....	36
2.2.7 fMRI attention task.....	38
2.2.8 fMRI analysis.....	38
2.3 Results.....	41
2.3.1 Localiser.....	41
2.3.2 Predictability.....	42
2.3.3 Adaptation.....	43

2.3.4	View shift sensitivity .....	47
2.3.5	Behavioural measures.....	48
2.4	Discussion.....	51
2.4.1	Representation of scenes in place selective cortex.....	51
2.4.2	Non-place selective VOT regions .....	52
2.4.3	The effects of low level visual change .....	53
2.4.4	Links between topographical memory and brain response .....	54
2.4.5	View sensitivity outside the VOT .....	56
2.4.6	Summary and Conclusions .....	58
<b>3</b>	<b>Processing of objects in object selective cortex: effects of parametric manipulation of viewpoint.....</b>	<b>60</b>
3.1	Introduction .....	60
3.1.1	Building invariant representations of objects.....	60
3.1.2	Visual object processing in the brain .....	62
3.1.3	View sensitivity throughout the VOT.....	63
3.1.4	Summary and aims.....	65
3.2	Methods .....	67
3.2.1	Participants .....	67
3.2.2	Imaging parameters .....	67
3.2.3	Object stimuli .....	67
3.2.4	Behavioural task.....	68
3.2.5	Localiser scan .....	68
3.2.6	Adaptation scan .....	69
3.2.7	fMRI attention task .....	70
3.2.8	fMRI analysis .....	71
3.3	Results .....	74
3.3.1	Localiser.....	74
3.3.2	Scrambled control maintaining visual change.....	75
3.3.3	Adaptation.....	76
3.3.4	View shift sensitivity .....	78
3.3.5	Behavioural measures.....	80
3.4	Discussion.....	82
3.4.1	The effects of low level visual change .....	83
3.4.2	Links between object recognition memory and brain response .....	84
3.4.3	Effects of attention .....	85
3.4.4	Summary and Conclusions .....	86
<b>4</b>	<b>Parametric adaptation for size change in object selective cortex.....</b>	<b>88</b>
4.1	Introduction .....	88
4.1.1	Size invariance in object selective cortex .....	89

4.1.2	Summary and aims.....	92
4.2	Methods .....	94
4.2.1	Participants .....	94
4.2.2	Imaging parameters .....	94
4.2.3	Object stimuli .....	94
4.2.4	Adaptation scan .....	95
4.2.5	Localiser scan .....	99
4.2.6	fMRI analysis .....	100
4.3	Results .....	102
4.3.1	Localiser.....	102
4.3.2	Size sensitivity in object selective cortex.....	103
4.3.3	Whole brain size sensitivity.....	106
4.4	Discussion.....	108
4.4.1	LOC as an intermediate stage in object representation .....	109
4.4.2	Adaptation in retinotopic cortex.....	111
4.4.3	Summary and conclusions .....	113
<b>5</b>	<b>Mapping invariance in category selective cortex.....</b>	<b>115</b>
5.1	Introduction .....	115
5.1.1	Mapping without statistical thresholds .....	116
5.1.2	Selectivity and Invariance in the VOT .....	118
5.1.3	Spatial comparisons of invariance .....	120
5.2	Methods .....	122
5.2.1	Participants and Data acquisition .....	122
5.2.2	Experimental design and stimuli.....	122
5.2.3	Measures of invariance and selectivity.....	123
5.2.4	Mapping visualisation techniques .....	124
5.2.5	Invariance x selectivity colour maps .....	125
5.2.6	Vector colour mapping.....	126
5.2.7	Spatial analysis .....	127
5.3	Results 1: Object and scene view invariance.....	130
5.3.1	The relationship of object and scene processing.....	134
5.3.2	Spatial analysis: selectivity .....	137
5.3.3	Spatial analysis: view invariance .....	139
5.3.4	Discussion 1: view invariance in the ventral stream.....	141
5.4	Results 2: Object Size Invariance .....	143
5.4.1	Discrete baseline measure of size invariance .....	144
5.4.2	Spatial analysis: size sensitivity x identity.....	147
5.4.3	Spatial analysis: discrete baseline invariance measure .....	149
5.4.4	Discussion 2: size invariance in the ventral stream .....	151

5.5	Results 3: Object size and view sensitivity.....	152
5.5.1	Spatial analysis .....	154
5.5.2	Discussion 3: separate forms of object invariance .....	156
5.6	Discussion.....	158
5.6.1	Statistically assessing the reliability of spatial maps .....	159
5.6.2	Large scale organisation of the ventral stream .....	160
<b>6</b>	<b>Summary and general discussion.....</b>	<b>162</b>
6.1	Hidden patterns of VOT organisation .....	166
6.2	Conclusions .....	167
	<b>Appendix A.....</b>	<b>168</b>
	<b>Appendix B.....</b>	<b>170</b>
	<b>Appendix C.....</b>	<b>172</b>
	<b>Appendix D.....</b>	<b>175</b>
	Within subject, across experiment analysis.....	175
	Spatial analysis .....	180
	Discussion: the robustness of spatial maps of invariance in individuals.....	182
	<b>References .....</b>	<b>183</b>

## List of figures

<b>Figure 1.1.</b> Organization of the ventral visual pathway.....	2
<b>Figure 1.2.</b> Schematic layout of human visual areas presented on an unfolded brain. .....	4
<b>Figure 1.3.</b> The logic of fMRI-adaptation to infer selectivity or invariance.....	15
<b>Figure 1.4.</b> Models of fMR-adaptation and underlying neuronal response.....	17
<b>Figure 1.5.</b> Hypothetical stimulus space, in which all possible objects could be placed.....	19
<b>Figure 1.6.</b> Parametrically increasing view shift blocks in representation space.....	20
<b>Figure 2.1.</b> Examples of conditions in the adaptation scan.....	37
<b>Figure 2.2.</b> Category selective, localiser scan ROI masks shown in MNI space.....	42
<b>Figure 2.3.</b> Response to all view shift conditions in all ROIs.....	44
<b>Figure 2.4.</b> Same place adaptation. ....	46
<b>Figure 2.5.</b> Linear responses to view shift. ....	48
<b>Figure 2.6.</b> Scatterplots of brain response and behaviour. ....	49
<b>Figure 3.1.</b> Examples of images used in the adaptation scan.....	70
<b>Figure 3.2.</b> Category selective, localiser scan ROI masks shown in MNI space.....	74
<b>Figure 3.3.</b> ROI responses to intact and scrambled 10° shift conditions.....	76
<b>Figure 3.4.</b> Response to all view shift conditions in all ROIs.....	77
<b>Figure 3.5.</b> Whole brain same object adaptation. ....	78
<b>Figure 3.6.</b> Whole brain linear responses to view shift. ....	79
<b>Figure 3.7.</b> Whole brain correlations with object recognition memory.....	80
<b>Figure 3.8.</b> LOC ROI correlation with object recognition memory. ....	81
<b>Figure 4.1.</b> Hypothetical responses for parametrically size shifting object stimuli...	93
<b>Figure 4.2.</b> Selection of object images used in the adaptation experiment.....	95
<b>Figure 4.3.</b> Low & high size change examples (fixed mean size $8.5^\circ$ ). ....	97
<b>Figure 4.4.</b> Object image silhouettes. ....	99
<b>Figure 4.5.</b> Category selective, localiser scan ROI masks shown in MNI space.....	102
<b>Figure 4.6</b> BOLD responses for same and different identity objects in all ROI's.....	104
<b>Figure 4.7.</b> Adaptation across size shift conditions in all three ROIs.....	106
<b>Figure 4.8.</b> Whole brain adaptation and linear trends. ....	107
<b>Figure 5.1.</b> Invariance x selectivity colour map. ....	126
<b>Figure 5.2.</b> Vector colour map. ....	127
<b>Figure 5.3.</b> Cross sections of MNI brain spatial sampling grid.....	129
<b>Figure 5.4.</b> Colour map of view invariance in object selective cortex.....	132
<b>Figure 5.5.</b> Colour map of view invariance in place selective cortex. ....	133

<b>Figure 5.6.</b> Vector colour maps of selectivity and view sensitivity.....	136
<b>Figure 5.7.</b> Place and object selectivity surface plots.....	138
<b>Figure 5.8.</b> Landscape and object view-invariance surface plots.....	140
<b>Figure 5.9.</b> Colour maps of size invariance in object selective cortex.....	144
<b>Figure 5.10.</b> Colour map of discrete baseline size invariance in object selective cortex.....	146
<b>Figure 5.11.</b> Different and same object size-invariance surface plots.....	148
<b>Figure 5.12.</b> Object size-invariance surface plots (discrete baseline).....	150
<b>Figure 5.13.</b> Vector colour map of object size and view sensitivity.....	153
<b>Figure 5.14.</b> Object size- and view-invariance surface plots.....	155
<b>Figure D.1.</b> Vector colour maps of category selectivity for four individual subjects transformed to MNI space (FLIRT).....	176
<b>Figure D.2.</b> Colour maps of landscape view invariance in place selective cortex for four individual transformed to MNI space (FLIRT).....	177
<b>Figure D.3.</b> Colour maps of object view invariance in object selective cortex for four individuals transformed to MNI space (FLIRT).....	178
<b>Figure D.4.</b> Colour maps of same object size invariance in object selective cortex for four individuals transformed to MNI space (FLIRT).....	179
<b>Figure D.5.</b> Landscape view-, object view-, and object size-invariance surface plots.....	181



## List of tables

<b>Table 2.1.</b> MNI coordinates of regions of interest peaks .....	42
<b>Table 2.2.</b> Different-Same adaptation contrasts, all ROIs. ....	45
<b>Table 2.3.</b> Adaptation contrasts for each view shift condition and all ROIs. ....	45
<b>Table 3.1.</b> MNI coordinates of regions of interest peaks .....	75
<b>Table 3.2.</b> Different-Same adaptation contrasts, all ROIs. ....	78
<b>Table 4.1.</b> MNI coordinates of regions of interest peaks .....	103
<b>Table 5.1.</b> Selectivity measures spatial grid ANOVA.....	139
<b>Table 5.2.</b> Landscape and object invariance spatial grid ANOVA.....	141
<b>Table 5.3.</b> Different and same object size invariance spatial grid ANOVA.....	149
<b>Table 5.4.</b> Object adaptation size invariance spatial grid ANOVA.....	150
<b>Table 5.5.</b> Object view- and size-invariance spatial grid ANOVA .....	156
<b>Table D.1.</b> <i>Same identity invariance spatial grid ANOVA</i> .....	180

## Acknowledgements

I would like to express my deepest gratitude to my supervisor, Dr Tom Hartley, for his patience, generosity and support. Since I first arrived in York to start my MSc I wondered when I was going to be recognised as a complete fraud and asked to leave the building. Tom was in the best position to realise this but somehow refrained from doing so. He has shared with me an unparalleled enthusiasm for cognitive neuroscience, taught me when to trust myself and when to doubt myself, and above all, never failed to keep me honest.

Thank you to my research committee for providing useful advice and feedback, and for keeping me on track; the ESRC for funding my research; and to all the people in the Department of Psychology and the York Neuroimaging Centre for providing a helpful and supportive working environment all these years. In particular, my closest friends from my time here in York, Shane Lindsay and Alex Reid without whom I don't think I could have clung to sanity long enough to finish this thing.

Thanks to my beautiful and understanding girlfriend Harriet, who has supported me, and miraculously not given up on me to the brink of madness and back again. I'm sorry I never quite managed to figure out the secret to a healthy work life balance.

To my family, Mum and Dad, thank you for believing in me, and giving me your unwavering love and support. Thank you for making me the person I am, and being proud of me no matter what.

## Declaration

This thesis comprises the candidate's own original work and has not previously been submitted to this or any other University for a degree. Selected aspects of the research described in this thesis have been presented elsewhere.

The data for experimental chapter 2 were gathered as part of group project under the supervision of the candidate. This experiment has been independently written and submitted for the completion of an MSc in Cognitive Neuroscience at the University of York by Sumyah Alnajashi, Samyogita Hardikar and Anastasia Pavlidou.

Experimental chapters 2 and 4 were presented as posters at the Organisation for Human Brain Mapping 2010 and Society for Neuroscience 2011 respectively.

Racey, C., Alnajashi, S., Hardikar, S., Pavlidou, A., & Hartley, T. (2010). Processing of landscapes in place selective cortex: effects of parametric manipulation of viewpoint. *Proceedings of the Organization for Human Brain Mapping, Neuroimage*, 51.

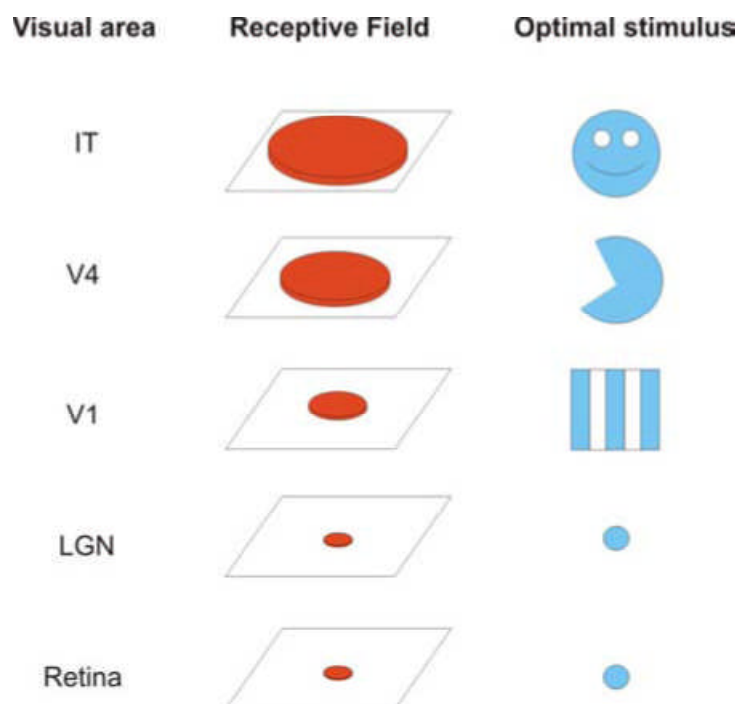
Racey, C., Andrews, T. J., Morland, A. B., & Hartley, T. (2011). Parametric adaptation for size change in object selective cortex. *Soc Neuroscience Abstracts*, 484.09.

## **1 The ventral occipital temporal pathway**

Everyday recognition of items in the world appears to be a relatively effortless process. It is something we do many times from moment to moment, even when faced with extremely brief glimpses and partial or degraded views of items. The ease with which we do such everyday tasks disguises the vast complexity that must underlie any computational system capable of carrying them out. The visual system of the brain has been extensively studied, and significant progress has been made in understanding the mechanisms of low level vision. However, as visual processing progresses from low-level visual regions towards regions specialised for memory, there remains debate about factors that govern functional organisation. There are brain regions which seem to process particular categories of stimuli, such as faces or landscapes. Yet there are suggestions that these regions may be better thought of as comprising a multi-dimensional representation space forming a continuous map, or multiple overlapping maps, the properties of which are not yet understood. This introductory chapter reviews what is known about key brain regions in the ventral occipital temporal pathway, known to be involved in recognition processes. It then goes on to discuss several theoretical frameworks for the topographical organisation of these regions, and describes approaches, both theoretical and methodological, that can be used to further understand how brain regions involved in visual recognition operate.

The mechanisms underlying how the brain accomplishes recognition are not yet fully understood. However, it has been established from studies of the anatomical pathways in the macaque monkey, that information flows along the ventral occipital temporal pathway in a hierarchical fashion, the bottom of the hierarchy includes the retina and the lateral geniculate nucleus (LGN), with the hierarchy culminating in the entorhinal cortex and the hippocampus (Felleman & Van Essen, 1991).

There are two distinct properties which are thought to encapsulate how information is transformed throughout the ventral stream, these are, ‘selectivity’ and ‘invariance’, thought to gradually increase throughout the visual hierarchy (Ison & Quiroga, 2008; Rust & DiCarlo, 2010). Selectivity refers to the fact that at each successive stage of the visual hierarchy, the response profile of the region, or the individual cells within it, is thought to be driven by more and more complex, yet specific, conjunctions of features (see **Figure 1.1**). These increasingly selective responses may ultimately culminate in representations of particular object categories or even specific exemplars to be recognised. Invariance refers to the idea that it is necessary that responses become increasingly tolerant to changes in the visual input, such as changes in position or size in the visual field, or changes in viewpoint or occlusion, that would otherwise make recognition impossible. The role of the ventral stream can be thought of as establishing representations that are increasingly *selective*, while also increasingly *invariant* to visual change.



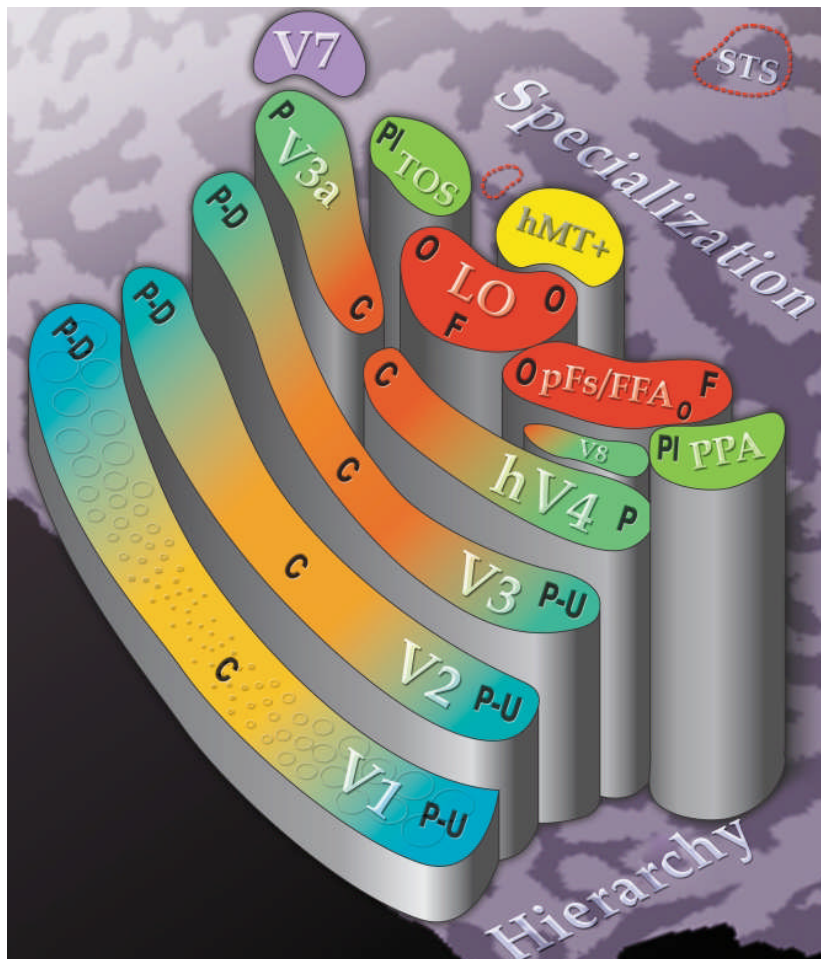
**Figure 1.1.** Organization of the ventral visual pathway. Along the ventral stream, the receptive fields of cells increase and the optimal stimulus becomes progressively complex (from Ison & Quiroga, 2008).

There is evidence for such gradual increases in both selectivity and invariance progressing through the visual processing hierarchy. Throughout the successive stages of the ventral stream, neurons respond to increasingly complex image features, and appear to do this by combining the features encoded by neurons at earlier stages. Areas V2, V4 and posterior inferotemporal cortex (IT) respectively, show increasing responsiveness for stimuli of increasing complexity (Brincat & Connor, 2004; Gallant, Braun, & Van Essen, 1993; Pasupathy & Connor, 1999). Evidence for gradual increases in invariance to changes in position and scale are also suggested by the gradual increases in receptive field size along the ventral stream (Kobatake & Tanaka, 1994; Rolls, 1994).

The increased complexity of the response properties of successive stages in the visual processing hierarchy, also gives rise to increased difficulty in building a full understanding of processing. Early visual regions such as primary visual cortex (V1) are amongst the most well understood regions in the brain. A great deal is known about the function of V1 at various different spatial scales, from the response properties of individual cortical columns (Hubel, Wiesel, & Levay, 1977), to the clear relationship between the large scale topographic organisation of V1, and layout of the photoreceptors in the retina (retinotopy; Holmes, 1945; Horton & Hoyt, 1991). The representation of the visual world found in V1 is best described by a two-dimensional, retinotopic, coordinate frame that can be referenced relative to the centre of the visual field. As one progresses along the visual processing hierarchy the representation moves away from this coordinate frame and becomes increasingly abstract (Ungerleider & Mishkin, 1982; Van Essen & Maunsell, 1983). Anterior regions show increasing selectivity for particular properties of visual stimuli and at the most anterior end of the ventral stream, regions tend to show selectivity for high level, abstract features that can be more related to perceptual properties of the stimulus than to the low level sensory properties (Sheinberg & Logothetis, 1997). To further complicate the issue, many extrastriate visual regions continue to show some form of retinotopic organisation in addition to particular kinds of feature selectivity (Larsson & Heeger, 2006), suggesting that at the large scale

population level, multiple types of information are represented at different spatial scales.

Grill-Spector and Malach (2004) published a large scale review of findings and methods employed to uncover the functional properties of the human visual processing hierarchy. They summarise multiple findings from functional neuroimaging and electrophysiology, and describe how two orthogonal principles, functional specialization, and hierarchical processing, can be seen laid out on the unfolded human visual cortex (see **Figure 1.2**). Perception and recognition are performed by neurons in a hierarchical network progressing from the mostly visually responsive V1, through to the medial temporal lobes (MTL), implicated in memory processes (e.g., Squire, Stark, & Clark, 2004).



**Figure 1.2.** Schematic layout of human visual areas presented on an unfolded brain. This figure illustrates the orthogonal axes of hierarchy and specialization. The visual areas are arranged in a staircase fashion to illustrate the hierarchical sequence of

*increased abstraction leading from primary visual cortex to high-order visual areas. The orthogonal, specialization axis is illustrated through the colour scale. The specialization is manifested in early cortex as a transition from central (C) to peripheral (P) visual-field representations. In higher-level cortex, the specialization is manifested as a transition from regions that respond preferentially to objects/faces (O, F), and are related to central-biased, high-magnification representations, to regions that respond more strongly to places, buildings, and scenes, and are related to peripheral-biased, lower-magnification representations (from Grill-Spector & Malach, 2004).*

### **1.1 Parcellating the ventral stream**

The primary focus of much of the present thesis is the ventral stream recognition pathway or extrastriate ventral visual cortex (VVC), and the functionally defined category selective regions that lie within it. This region has two main anatomical subdivisions. The first is the Ventral Occipital Temporal (VOT) cortex, which corresponds to the extrastriate ventral surface of the brain, including the parahippocampal and fusiform gyri and extending into the MTL (Grill-Spector & Malach, 2004; Tanaka, 1996). The second is the lateral occipital cortex (referring to the anatomical region as distinct from the functionally defined Lateral Occipital Complex (LOC), which overlaps with the VOT), this extrastriate region is located on the lateral bank of the fusiform gyrus and extends laterally and dorsally (Grill-Spector, Kourtzi, & Kanwisher, 2001).

At the level of the VOT, and what is often considered to be its monkey homologue, IT, the cells are primarily selective for high level stimulus properties. They are characterised by selectivity for particular categories, and can respond regardless of where in the visual field the stimulus appears. Several early monkey neurophysiology studies were able to show populations of cells with a greater response to one particular category than any other. For instance, Gross et al. (1972) reported cells in IT that were selective for a monkey's hand, and separate cells responding maximally for face stimuli. Another example is Desimone et al. (1984), who found cells in the Superior Temporal Sulcus (STS) of the IT region that



responded maximally to both monkey and human faces, but less well to body parts or objects, or to faces with the internal feature positions shuffled.

These findings in the monkey literature have been confirmed and greatly extended in the human neuroimaging literature. FMRI, measuring changes in blood oxygenation (BOLD), in particular has been used to address questions about high level visual processing. Several category selective regions have been identified and labelled. An area in the lateral occipital cortex has been found showing a greater BOLD response to objects compared with scrambled objects and textures, and termed the LOC (Malach et al., 1995). This region extends from the posterior inferotemporal sulcus into the posterior fusiform gyrus and there is evidence that the region can be subdivided at the border between these two anatomical features, the anterior subdivision showing more invariant adaptation than the posterior region (Kourtzi, Erb, Grodd, & Bulthoff, 2003).

Several other category selective regions have been identified within the anatomical area VOT. The Fusiform Face Area (FFA), a region which shows greater response to face than to place stimuli (Kanwisher, McDermott, & Chun, 1997) and the Parahippocampal Place Area (PPA), a region lying in the parahippocampal gyrus, which responds more to place stimuli than to face stimuli (Epstein & Kanwisher, 1998). In the case of both the PPA and the FFA, scrambled or otherwise reconfigured controls have been conducted showing that these regions respond more to their preferred category than to the low level properties present in the stimulus images (Andrews, Clarke, Pell, & Hartley, 2010).

Since the discovery of areas within the VOT that are selective for particular visual categories, it has become common to further subdivide the region, identifying additional areas that show selective responses for more specific aspects of the visual world. The Extrastriate Body Area (EBA), adjacent to, and often overlapping, the LOC, is a region which shows particular selectivity for bodies and body parts (Downing, Jiang, Shuman, & Kanwisher, 2001). An additional, spatially distinct region has been proposed, close to, and overlapping the FFA, the Fusiform Body Area (FBA) (Peelen & Downing, 2005; Schwarzlose, Baker, & Kanwisher, 2005).

It has been suggested that EBA and FBA can be functionally dissociated, with a more selective activation for local body parts in EBA relative to more holistic images of the human body in FBA (Taylor, Wiggett, & Downing, 2007).

Additionally, a region which shows considerable overlap with the FFA is the Visual Word Form Area (VWFA; Cohen, Jobert, Le Bihan, & Dehaene, 2004). This region shows a strong response to visually presented words or letter strings, and has sparked considerable debate as to whether the VWFA represents letter strings specifically or has a more general role in visual analysis (Cabeza & Nyberg, 2000; Reinke, Fernandes, Schwindt, O'Craven, & Grady, 2008; Starrfelt & Gerlach, 2007). The FFA and VWFA occupy much the same cortical space, although both have been shown to have hemispheric biases, with the FFA showing a tendency towards a right lateralised response (Cabeza & Nyberg, 2000). The VWFA on the other hand, shows a clear left lateralised activation (Cohen et al., 2002), suggestive that functional connections with language systems may influence the selectivity of the VOT.

The growing complexity of describing category selective subdivisions in the VOT has led to considerable debate as to whether it is meaningful to continue to parcellate category selective cortex into more and more specialised regions, or whether other factors give rise to the lawful, replicable patterns of selectivity that can be observed in the VOT (Op de Beeck, Haushofer, & Kanwisher, 2008). Limitations in conventional fMRI approaches may mean that subtle patterns occurring below the resolution of current scanning techniques may be undetectable. Or meaningful information may be contained in the distributed pattern of activity across all of category selective cortex such that univariate methods (looking at each voxel individually) may be unable to fully describe it.

## **1.2 Organising principles of the VOT**

Whilst it is clear that the VOT region contains populations of neurons which show selective responses for particular kinds of stimuli, it is not clear what the ontology of these patterns of functional activity might be, and what might govern the

organisation of this high level visual area. Several theoretical accounts have been put forward as to how to best account for the category selectivity observed in the VOT.

Kanwisher and Downing et al., (2001; 2000) have taken the view that there are a limited number of discrete regions at the level of the VOT, which can be thought of as domain specific modules, specialised for the processing of particular categories of object. Central to this argument is the idea that regions such as the FFA are specialised only for faces, and as such, are not involved in the processing of other categories of stimuli. The overall suggestion is that the VOT is for general object recognition but contains several additional sub-regions for domain specific processing.

An alternative account has been put forward by Haxby et al. (2001), who found that the distributed pattern of BOLD activity in the VOT region was sufficient to classify the category of stimulus being viewed by the participants. This held true even when the activity within the key functionally defined regions was removed from the analysis. It was therefore suggested that the VOT region contains an 'object form topography', in which the representation of an object is reflected by a distinct pattern of response across all of VOT, and this distributed representation is responsible for visual perception. Haxby's account in its extreme form is also problematic. It has been criticised because similar classification methods to those used by Haxby found that while it was possible to classify preferred vs. non-preferred stimuli types using activity from functionally defined regions, it was not possible to classify between two different non-preferred stimuli types (Spiridon & Kanwisher, 2002).

The particular, and extreme deficits for face processing, encountered by prosopagnosic patients who have suffered focal damage to the fusiform region, seem to suggest a particular functional specificity for the FFA (Hecaen & Angelergues, 1962; McNeil & Warrington, 1993). A key example is a case report by Wada and Yamamoto (2001), of a man with an unusually restricted lesion within the right FFA. This patient was able to discriminate a face from a non-face, but was

critically impaired when discriminating between particular faces. This damage did not result in any other impairments, suggesting that the region is non-critical for processing other kinds of visual stimuli.

An alternative framework was put forward by Tarr and Gauthier (2000), who suggest that the pattern of activity in the VOT region in response to particular stimulus categories, is driven not by category selective neuronal populations, but by neurons specialised for particular kinds of processing, which vary depending on the kind of stimuli viewed and the kinds of task being carried out. In the case of face processing, recognition processes always involve discerning identity on the basis of subtle variation of particular features and their configuration. And faces are typically viewed from a limited number of viewing angles. Object processing on the other hand, could involve large differences in any number of featural or configural aspects of the object, which are far less constrained than faces tend to be. They argue that VOT specialisation, and indeed all kinds specialisation in the brain, is on the basis of 'process' rather than domain specificity. This account predicts that regions such as the FFA, which can be thought of as performing expert within-category discrimination, should show equivalent responses to other object categories for which similar kinds of within-category processing is extremely well learned.

This idea has received some empirical support; Gauthier, Skudlarski, Gore, and Anderson (2000) found that in bird and car experts, the region activated when viewing the participants own expert stimuli showed substantial overlap with the FFA. Others who fail to find similar effects argue that the different findings arise from methodological and design differences, and that, in the case of face processing, the neural specificity observed is a 'special case' of object recognition with specialised face selective neuronal populations (McKone & Robbins, 2007).

A further factor that may govern the organisation of the VOT region is the basic retinotopic organisation present in preceding visual areas. Levy et al. (2001) identified a large scale centre-periphery organisation in the topography of functionally defined category selective regions, with the foveal part of the map

represented in the lateral parts of the VOT, and the peripheral part mapping onto medial regions. Levy et al. suggest that this is driven by the position in the visual field in which each regions preferred stimuli typically appears, viewed landscapes generally encompass the entire visual field, and so are represented in the peripheral/medial part of the map, whereas faces tend to be foveated, and as such occupy a central/lateral area.

A recent review by Op de Beeck et al. (2008) attempted to reconcile these competing theoretical ideas about the organisation of the VOT. They suggested that the functional mapping architecture that can be seen in early retinotopic cortex, where maps of all orientations of line, and ocular dominance, are embedded in a columnar structure within a retinotopic map of the full visual field, may also apply to high level areas such as the VOT. Several organising principles for the VOT region have been suggested, such as object category, functional connectivity, process, and eccentricity. Op de Beeck et al. suggest that these may not be mutually exclusive, and high level visual cortex may contain overlapping or embedded maps encoding multiple stimulus dimensions. The nonlinear summation of activation in these maps could give rise to the pattern of category selectivity that is widely observed in fMRI studies, and depending on the experimental design and analysis method, give the impression of either domain specific modules or highly distributed general processing structures.

In early visual cortex the visual field is mapped onto the cortical surface, and therefore the key large scale organising principle is spatial position. It is possible to map the organisation of V1 by continuously varying visual field position of high contrast chequerboard stimuli (Engel et al., 1994; Sereno et al., 1995). A recent study by Goeseart and Op de Beeck (2010) attempted to apply the logic of retinotopic mapping to category stimuli by gradually morphing their stimuli between faces and houses. Their aim was to look for travelling waves of activity that were in phase with the change in their stimuli, and potentially uncover a large scale mapping throughout VOT. They were however, unable to find any meaningful activity spanning the PPA, FFA and intermediate cortex. Furthermore, the activity

within the category selective regions showed no clear responses to the intermediate morph stimuli. A similar experiment by Tootell et al (2008) using a blocked design approach instead of a phase-encoding method, also found no evidence of a topographic map spanning category selective cortex.

These findings highlight the difficulties in parameterising high level visual stimuli. Nonlinearities in the way the activation of multiple overlapping maps are summated create difficulties in determining an appropriate manipulation that can continuously activate the full extent of the VOT. It remains an open question whether high level stimulus manipulations exist that can be used to link category selective regions as part of larger scale maps (Op de Beeck et al., 2008).

### ***1.3 Invariance as an organising principle***

As argued above, a key feature of the visual system is that it abstracts away from the low level input representation and builds high level representations that facilitate invariant recognition. It is also very likely that this process takes place gradually, and in stages (Felleman & Van Essen, 1991; Grill-Spector & Malach, 2004). Different kinds of visual stimuli that humans routinely encode, store, and recognise, may be processed in different ways. For example, when reading text it is unlikely that letter strings will be seen from an alternate view point, and in many cases the meaning of letters is critically dependent on their exact configuration. In this case it is possible that the system does not need to build representations that support rapid recognition from any view point. Conversely, it may be the case that a representation of landscapes needs to be computed that is invariant to changes in view point because in the course of spatial navigation, it is often necessary to recognise previously unseen views of an environment (Milner & Goodale, 1995; Presson, DeLange, & Hazelrigg, 1989).

The nature of the representation, and the way in which it is established may vary depending the stimulus being processed, and the properties of the stimulus that are meaningful for recognition (Gauthier et al., 2000). The brain may need to

establish different kinds of invariance for different properties, and the degree of importance for each of these properties may vary. It is therefore suggested, that the topographical organisation of the VOT region may be influenced by the forms of invariance required for different stimuli and different processing tasks. The present thesis develops a paradigm whereby particular types of invariance can be quantified. Combining this with measures of neuronal selectivity it is possible to develop new ways to characterise how selectivity and invariance change throughout category selective cortex (see chapter 5 introduction).

#### **1.4 Using fMRI to investigate high level representations**

A great deal of what is known about neuronal representations in the ventral stream comes from monkey electrophysiology, in which the response properties of individual neurons are probed under various stimulus viewing conditions. However, high level visual representations are likely to be encoded over populations of neurons where the response of individual cells may fail to show the whole picture. The level of description of fMRI seems ideally suited to the study of population representations, and indeed the representation of the visual field in early visual regions is well characterised by retinotopic mapping techniques using fMRI (Engel et al., 1994; Sereno et al., 1995; Sereno, Pitzalis, & Martinez, 2001).

In primary visual cortex, embedded within the retinotopic map, are maps of ocular dominance and orientation selectivity. Representations of all possible orientations of line exist within a single square millimetre patch of V1 (Bartfeld & Grinvald, 1992). This is an example of maps of feature selectivity embedded within a larger topographic map, and yet these maps of ocular dominance and orientation occur below the spatial resolution of conventional fMRI, and should be undetectable using such methods. There are however, techniques that have been developed in recent years that have proven effective in detecting responses that occur at a sub-voxel spatial scale.

### 1.4.1 MVPA

Multi-Voxel Pattern Analysis (MVPA) is a technique that is able to detect weak signal biases that would not be detectable using conventional univariate statistics (Haxby et al., 2001; Norman, Polyn, Detre, & Haxby, 2006). Kamitani and Tong (2005) were able to demonstrate the power of the technique by performing accurate classification of the orientation content of viewed stimuli using the BOLD signal from within primary visual cortex, indicating that it is possible to detect the activity of sub-voxel level selectivity in the absence of global selectivity.

The MVPA technique has been used to decode weak signal biases resulting from the activation of populations of cells that are too small to detect with standard fMRI, but are none the less well understood from single unit work. As discussed above, the nature of the representations in the VOT may be the result of multiple overlapping maps that sum non-linearly to give rise to the patterns of activity typically described by univariate methods (Op de Beeck et al., 2008). It is possible to use MVPA to detect biases that take place at any spatial scale and in many cases without being able to describe the nature of the underlying representation, for example to classify particular exemplars of visual stimuli or even reconstruct the contents of perception using the signal in the VOT cortex (Kriegeskorte, Formisano, Sorger, & Goebel, 2007; Nishimoto et al., 2011).

Recent attempts have been made to adapt these techniques to produce a common, high-dimensional description of the representation space of the VOT (Connolly et al., 2012; Haxby et al., 2011). The logic of this approach is to co-register the brains of individual participants, not on the basis of the spatial structure of the brain anatomy, but on the basis of the parts of the brain that show common correspondence, across individuals, during complex visual perception. This 'hyperlignment' method uses the brain response patterns from viewing a full-length movie, and the resulting coherence in patterns of BOLD activity over time, to transform individual brain activation data into a common, high dimensional space. Haxby et al.'s approach is able to achieve excellent classification performance between subjects using this new method, where previously only effective within-



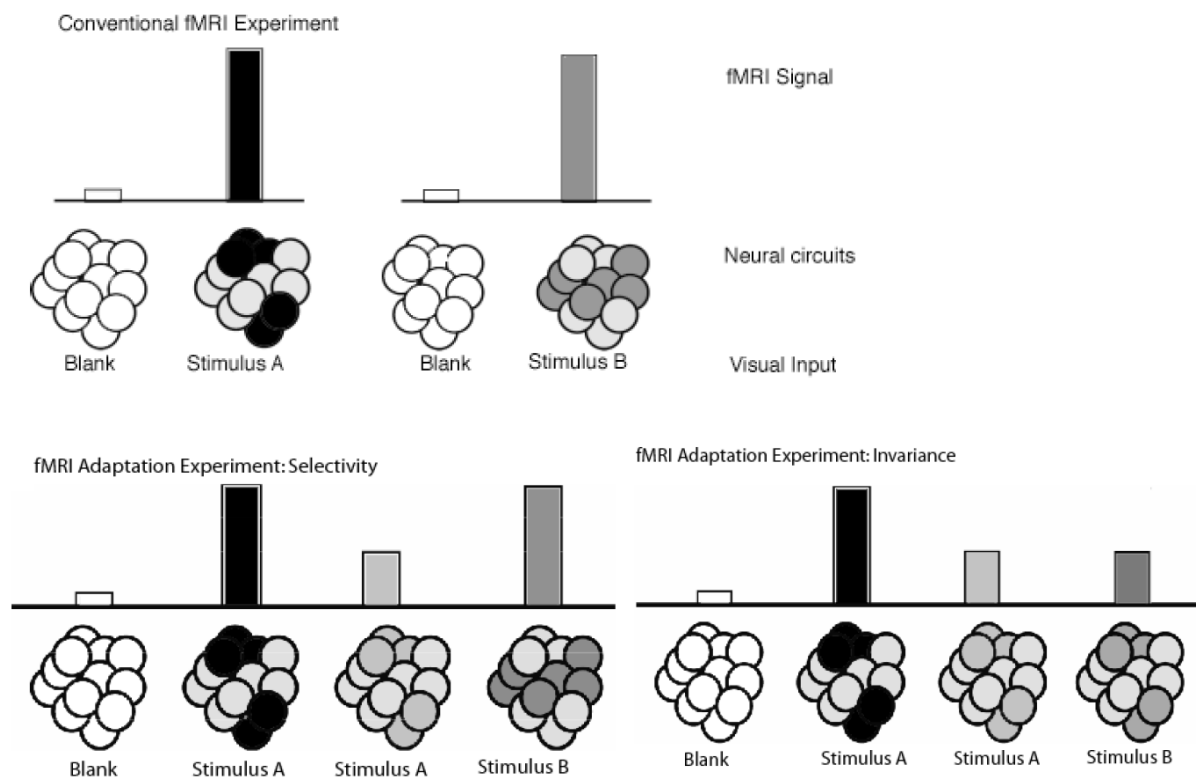
subject classification was possible. The resulting multidimensional space, constrained with principal components analysis, is thought to describe the space used within the visual system, to represent complex visual stimuli (Haxby et al., 2011). A limitation of this approach however, is that it remains difficult to understand exactly what individual dimensions within such a putative high-dimensional representation space correspond to in the visual world.

#### 1.4.2 *fMRI-adaptation*

A further method that has been used to overcome the limited spatial resolution of fMRI relies on the phenomenon of adaptation. This has long been known about in vision science and used as a tool to study low level processing mechanisms (Harris, 1980). For example, repeatedly fixating on lines or gratings of a particular orientation will change the subsequent perception of lines of a different orientation (Gibson & Radner, 1937). Such perceptual phenomena suggest that some internal brain activity must change with repeated exposure; in this case, it is thought to be the orientation selective neurons in primary visual cortex.

In fMRI, adaptation has been used to characterise the functional properties of neurons at subvoxel resolutions (fMR-a; Grill-Spector & Malach, 2001; Weigelt, Muckli, & Kohler, 2008). Presenting repeated, identical object images in fMRI will result in attenuation of the signal, and depending on whether that attenuation is robust to changes in subsequent images, or sensitive to them, can be used to make inferences about the nature of the underlying representation (**Figure 1.3**). Although measurements are taken at the resolution of conventional fMRI, the attenuation of signal could be driven by a reduction in activation of sub-populations of neurons within a single voxel. This is particularly useful for the study of representation, as populations of neurons representing different aspects of the stimulus could co-exist within the same voxel (Kourtzi & Grill-Spector, 2005). Fang, Murray, Kersten, & He (2005) were able to show that it is possible to reliably detect orientation dependent adaptation in early visual cortical regions using fMR-a.

Several single unit recording studies have tested for correspondence between neural adaptation and fMR-adaptation effects across species, and across the different kinds of experimental design (Krekelberg, Boynton, & Van Wezel, 2006; Sawamura, Georgieva, Vogels, Vanduffel, & Orban, 2005; Sawamura, Orban, & Vogels, 2006). Although there are caveats that the relationship is complex and less straightforwardly interpreted than is often assumed, it is generally thought that fMR-a is a reliable measure of neuronal adaptation, and that there is good correspondence between humans and monkeys in adaptation effects (Krekelberg et al., 2006; Weigelt et al., 2008).



**Figure 1.3.** *The logic of fMRI-adaptation to infer selectivity or invariance. The upper graph describes how conventional fMRI is used to measure neural responses to particular stimulus types. When the same stimulus is repeated the cells responding to it show an attenuation of signal. The lower left graph shows how a release from this signal attenuation accompanying a change in the stimulus can be used to infer selectivity. The lower right graph shows how a sustained signal attenuation accompanying a change in the stimulus can be used to infer invariance (from Kourtzi & Grill-Spector, 2005).*

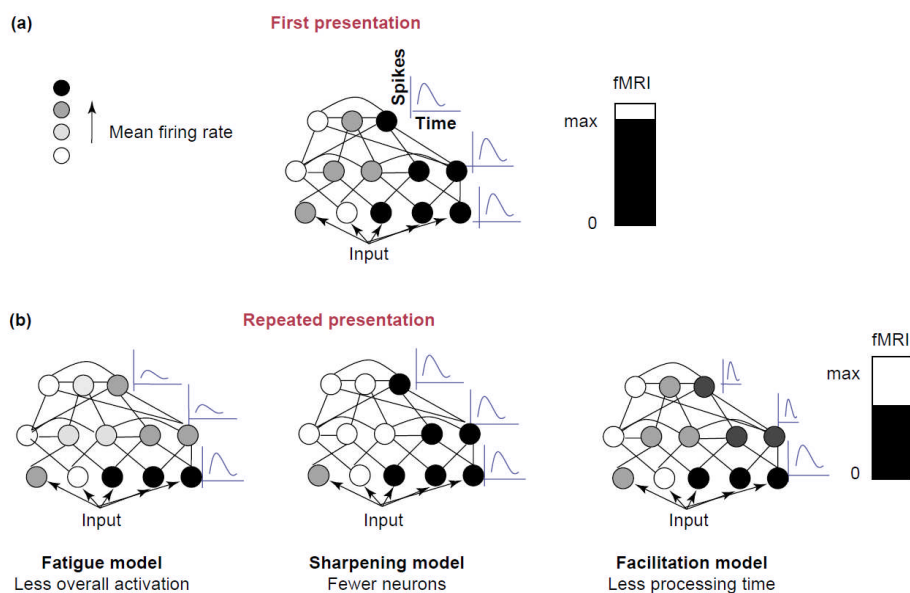
Three models have been suggested for how fMR-adaptation relates to underlying neuronal response (**Figure 1.4**). (1) The *Fatigue model* in which the neurons that are responding to the stimulus show a decrease in amplitude to subsequent presentations; (2) the *Sharpening model*, whereby, the activated neurons become more tuned to the repeated stimulus and fewer neurons become involved in the representation with subsequent presentations; finally, (3) the *Facilitation model*, whereby residual activation allows for repetitions of the signal to more easily activate the same neurons again, resulting in the representation being active for less time. Each of these models reflects a gradual increase in processing efficiency, and all of them result in similar patterns of BOLD response in fMRI, which aggregates signal over both space (e.g. a single 3mm voxel) and time (the three seconds needed to acquire a functional volume). Grill-Spector, Henson, & Martin (2006) have suggested subtle ways in which the different models might manifest themselves, and of course it is possible that these models are not mutually exclusive. Different kinds of adaptation may occur depending on cognitive task or location in the brain. Kohn and Movshon (2004), for example, found that sharpening best describes neural firing in MT, whereas fatigue best described V1 responses.

There are several different experimental design types that exploit the phenomenon of adaptation, and there is some indication that the underlying neuronal effect may differ depending upon the design type used (Weigelt et al., 2008). Broadly speaking, there are two main design distinctions within the range of fMR-adaptation experiments typically employed to study neural representation. Event related approaches are the most commonly used for such questions, and these involve rapid presentation of two consecutive individual stimuli, such as two consecutive object images, and testing for significant reduction in signal for the latter image relative to the former (Kourtzi & Kanwisher, 2000). The event related approach is thought to test short-lag adaptation effects. Another commonly used approach is a block adaptation design (Grill-Spector et al., 1999). This involves stimuli presented rapidly in 'blocks', and relies on a reduction in signal over the

course of the block when some property is held constant. Activity in this condition is tested relative to a separate condition where the stimulus property of interest is varied throughout the block.

Sawamura et al., (2006) suggest that event related designs may overestimate adaptation effects when testing neural populations that show selectivity for the test stimuli, and that this issue is not as prevalent in block design experiments. They found that adaptation was possible even when different items that drive selective responses are used, and this was not true for neurons that are not selective for the stimuli. Therefore, the amount of adaption a cell or neural population shows interacts with its selectivity.

It has proved possible to draw inferences about the nature of neural representations in cortical regions using fMR-a, without a complete understanding of its neural basis (Grill-Spector et al., 2006). The majority of studies looking at representations within category selective cortex make an implicit assumption that the *Fatigue model* is true (Ewbank, Schluppeck, & Andrews, 2005; Grill-Spector et al., 1999; Vuilleumier, Henson, Driver, & Dolan, 2002). Exactly which of these models provides the best description of fMR-a remains an open question, and may have subtle but meaningful effects on the interpretation of fMR-a findings.



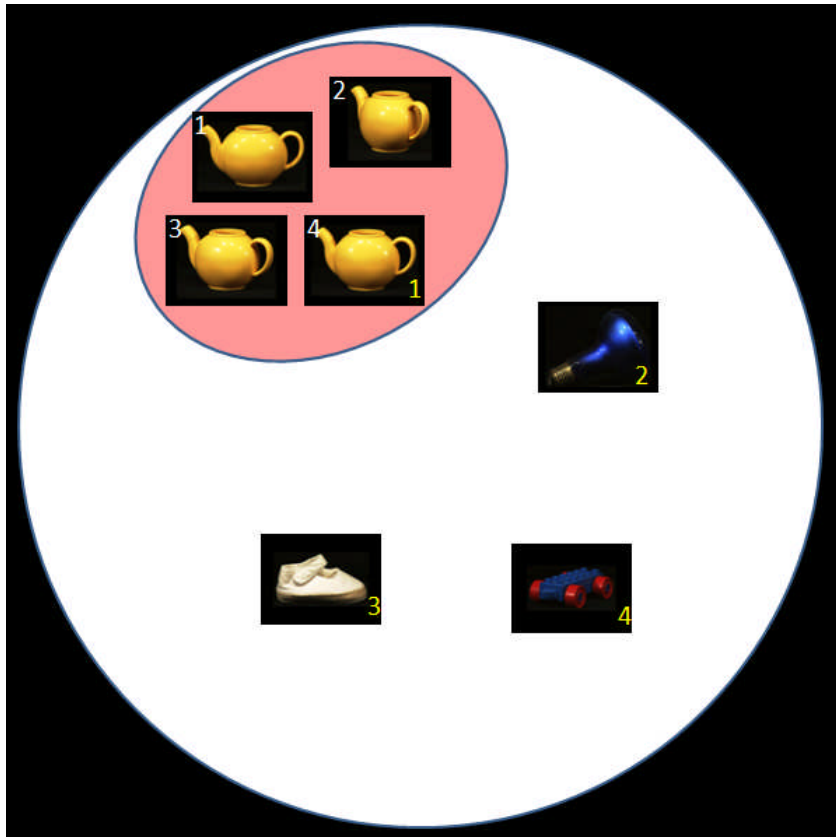
**Figure 1.4.** Models of fMRI-adaptation and underlying neuronal response.

(a) The visual stimulus is assumed to cause activity in early visual cortex before

*being processed in a hierarchical sequence of stages. The blue graphs indicate spiking of the neurons with highest response at each stage. (b) Because the BOLD signal integrates neuronal activity over time, all three of these models predict reduced BOLD for repeated stimuli, but for different reasons (from Grill-Spector, Henson, & Martin, 2006).*

### **1.5 Probing representational space**

A central aim of the present thesis is to investigate how the process of establishing invariance influences the organisation of the VOT and the regions within it. A paradigm is developed that uses fMR-adaptation as a tool to probe along potential dimensions within this representational space. Typically, in fMR-a experiments, when a high level property such as viewpoint is varied, maintaining identity, then sustained adaptation is used to infer whether or not the underlying neural representation is invariant to changes in view (Epstein, Graham, & Downing, 2003; Epstein & Higgins, 2007; Grill-Spector et al., 1999; Kourtzi et al., 2003; Vuilleumier et al., 2002). If there is recovery from adaptation then it is thought to reflect the neural population treating each view of the object as a new stimulus, and the representation is considered view-dependent. It might be unclear what aspects of the stimulus a given population of cells represents and fMR-a can be used to reveal when stimuli are close together in the representation space. **Figure 1.5** describes how conventional fMR-a experiments use same and different identity stimuli to sample within a putative representation space. Kourtzi & Kanwisher (2000) for example found similar levels of adaptation for both greyscale photographs of objects and equivalent line drawings of them. These images differed dramatically in terms of low level visual properties, but were similar in terms of shape and object identity. In this case it was possible to use this approach to investigate whether a given set of potentially very different stimuli are treated as close together, or far away by the extrastriate visual system.

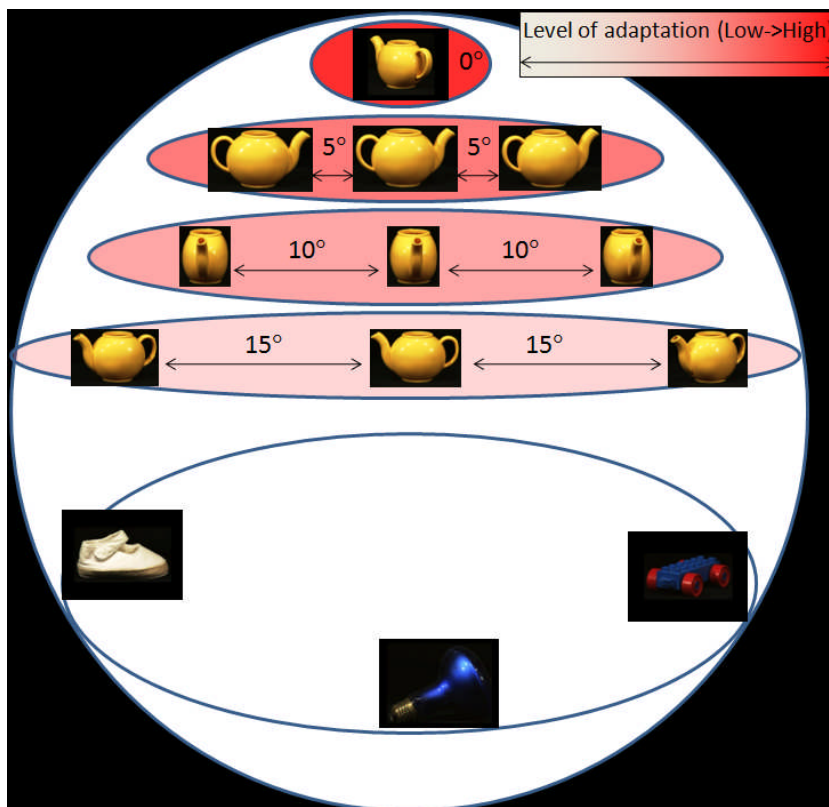


**Figure 1.5.** *Hypothetical stimulus space, in which all possible objects could be placed.*

*The number and nature of the dimensions used to describe this space are unknown. In order to drive adaptation, sequentially presented stimuli must be close together within this space. A block of four different objects (yellow numbers), are far apart in this space and as such do not drive adaptation. A block containing four teapots from different views (white numbers) show adaptation (pink ellipse) and it can be concluded that they are close together in representation space.*

FMR-a has proven to be a useful tool to study ventral stream representations, however, it is not without its limitations. The presence or absence of adaptation when a particular feature altered is used to infer that a representation is invariant, or not, to the manipulated property. A problem with this inferential step is that it assumes that invariance is an all-or-none phenomenon. There are several examples in the literature where different conclusions have been drawn about the nature of object or scene representations depending upon the degree to which the *same* condition was manipulated (see chapters 2 & 3 for detailed discussion).

The experiments within this thesis use a paradigm whereby the degree of change is parametrically manipulated across multiple conditions within a single experiment (see **Figure 1.6**). Change across multiple conditions increasing linearly can be thought of as sampling increasingly spread out distances in the representation space. It is common for fMR-a experiments to test for the presence or absence of adaptation, however, the present parametric adaptation paradigm assumes that the rate of change in neural adaptation across shift conditions describes the degree of sensitivity to the manipulated stimulus dimension. A measure of how steep the change in response is across multiple shift conditions, such as the slope of a linear fit, can be thought of as a measurement of sensitivity/invariance. A flat slope reflects invariance to a stimulus dimension a steep slope reflects sensitivity to it, with a continuum of possibilities in between.



**Figure 1.6.** Parametrically increasing view shift blocks in representation space. The amount of rotation between successive images within the block increases from  $0^\circ$ , in  $5^\circ$  steps, up to  $15^\circ$ . The distance in the representation space is increasing in a graded fashion. The measure of adaptation can be used to infer how close together items within a block are within this space.

## **1.6 Summary and aims**

In recent years it has been suggested that representation at the level of the VOT is more complex than was previously thought. The general picture that is beginning to emerge is that such representations are best thought of as a multidimensional space within which all visual stimuli can be located (Connolly et al., 2012; Dicarlo & Cox, 2007; Haxby et al., 2011; Op de Beeck et al., 2008). The response of a given region or voxel can be influenced by several stimulus dimensions at once. The nature of many of these dimensions is unknown or at the very least incompletely described. Understanding the topographical organisation and representational structure of the VOT remains a key challenge. Conventional fMRI approaches may be insufficient to describe such multidimensional complexity, and it will be necessary to extend them in order to be useful for such investigations.

Several extensions of classical fMRI have been developed, such as MVPA and fMR-adaptation, which allow for investigations into the sensitivity of a given region to particular stimulus dimensions. Any given stimulus has a location within the putative high dimensional space, and such methods can be used to test which potential dimensions play an important role in the representation in a given region or voxel. Varying a particular stimulus attribute can be thought of as spacing stimuli close together, or far apart, along a particular dimension. Whether or not multivariate patterns or adaptation are the same or different can allow inferences to be drawn about whether or not a particular region is invariant or sensitive to a dimension of interest. Such measurements of sensitivity and invariance may be independent of the selectivity of a given region.

The debate about the extent to which category selective cortex is organised into domain specific modules or continuous graded change in function has thus far been influenced heavily by the tools used to investigate it. If, as Op de Beeck (2008) suggests, the organisation of the region is the result of multiple, continuous, overlapping maps, then existing approaches which test hypotheses about the presence or absence of particular kinds of invariance would be unable to reveal



such an organisation. In order to properly investigate this issue, it is necessary to measure the degree of sensitivity to particular stimulus dimensions within category selective regions, and throughout the VOT.

The present thesis develops a novel parametric adaptation paradigm whereby high level properties of the stimulus are varied linearly within a single experiment. Adaptation is measured by degree rather than as an all or none phenomenon. Initially, chapters 2 and 3 apply a parametric adaptation paradigm to investigate sensitivity to 3D viewpoint in the representations of landscapes and then objects in category selective cortical regions. Chapter 4 develops the paradigm further by applying it to both same and different identity objects. Size sensitivity of objects is investigated while controlling for the extent of visual field stimulation.

All of the experimental chapters employ regions of interest (ROI) and whole brain analysis methods equivalent to those used throughout the fMRI literature to investigate similar questions. These experiments develop and refine a parametric adaptation paradigm which is able to examine whether invariance in category selective cortex is an all or none-phenomenon. Finally, chapter 5 brings together data from the previous experimental chapters in order to investigate the functional architecture of different forms of invariance as they are established by the ventral stream for different stimulus categories. Novel analysis techniques are developed and a whole brain mapping approach is described whereby measures of selectivity and sensitivity/invariance are defined at the voxel level. These are used to visualise and measure changes in invariance for different stimulus types within and between category selective VOT regions.

## **2 Processing of landscapes in place selective cortex: effects of parametric manipulation of viewpoint**

### **2.1 Introduction**

The ventral and dorsal visual pathways have classically been characterised as the 'what' and the 'where' pathways (Ungerleider & Mishkin, 1982). Within this framework, the ventral stream is thought to gradually extract information from the visual input in order to remember and recognise items in the world. Information such as size and position in the visual field are discarded as they are not useful for recognition, and increasingly invariant representations are established. The dorsal pathway leads from early visual cortex towards effector systems, and is thought to support visually guided action. Spatial information such as the size and position relative to the body is computed from the visual input. This allows for movements such as orientation of the head and eyes towards an object, or for reaching out and grasping of an object. The size and spatial coordinates of objects in the world are computed by this pathway.

Implicit in this framework, is the idea that the ventral stream and the processes of visual recognition discard spatial information contained in the visual signal, leaving spatial processing to the dorsal pathway. The endpoint of the ventral stream however, is the medial temporal lobes and the hippocampus. These systems have long been implicated in declarative memory processes (Scoville & Milner, 1957). Additionally there is good evidence that hippocampal structures are involved in spatial navigation (O'Keefe & Nadel, 1978). There remains an ongoing debate about the best way to account for processing in these regions, and whether spatial processing is best described as a special case of mnemonic processing, or episodic memory processing as a special case of spatial processing (Bird & Burgess, 2008; Squire et al., 2004). Regardless of the outcome of this debate, there is little doubt

that spatial information is processed by the hippocampal formation, and a great deal of the input to these regions originates from ventral stream regions such as the parahippocampal cortex. Anatomically the hippocampus and entorhinal cortex (the anterior part of the parahippocampal gyrus) comprise the hippocampal formation which is at the apex of the ventral stream (Amaral & Lavenex, 2007). It is likely that spatial information is not only preserved in the ventral stream, but explicitly processed, in order to allow for such spatial navigation and memory processes.

There is evidence that the parahippocampal gyrus selectively responds to navigationally relevant landmarks. Janzen and Turennout (2004) asked participants to view a route through a virtual environment then, during a recognition task in the scanner, showed them objects encountered along this route. Regions of the parahippocampal gyrus were subsequently more active for objects seen at decision points from the route, than objects seen at incidental locations.

Milner & Goodale (1995) described an alternative to the 'two-cortical visual systems' theory of Ungerleider and Mishkin. They suggest that there are two kinds of spatial processing which recruit different kinds of information from the visual signal. The dorsal stream computes very precise information about spatial location relative to the eye/head/body (egocentric coordinates) in order to allow distal control of manual prehension and other skilled visuomotor acts. The ventral stream computes the spatial relationship of objects relative to each other and the surrounding environment (allocentric coordinates). Such processes necessarily include information about the identity of visual objects, and allows for perception and memory of the external world. This theoretical framework predicts that allocentric spatial coding will not exist at any point along the dorsal stream, but is established by the ventral stream.

There is evidence that there are allocentric representations of the environment in the hippocampus, which receives its input from the ventral stream via the parahippocampal, entorhinal, and perirhinal cortices. Humans with hippocampal damage show deficits in spatial processing and navigation, as well as topographical agnosia (Scoville & Milner, 1957). Furthermore, topographical

agnosia is often seen in cases of parahippocampal cortical damage (Aguirre & D'Esposito, 1999). The clearest evidence for allocentric processing however, comes from the discovery of place cells in the hippocampus of the rat (Muller & Kubie, 1987; O'Keefe & Conway, 1978). Place cells are individual cells which increase their firing rate as the animal moves through particular locations within its environment. The location, or 'place field', that any given cell responds to remains constant regardless of sensory input. The visual input could be completely different depending on the direction the animal is facing, yet the place field of the cell remains constant. A population of place cells provides an allocentric representation of all possible locations within a given environment. At the end point of the ventral stream representations are essentially view invariant and used to construct a representation of location.

The activity of place cells is thought to be supported by input from cells in the entorhinal cortex known as grid cells, which show a regular spatial firing pattern corresponding to an evenly spaced triangular grid throughout the environment (Hafting, Fyhn, Molden, Moser, & Moser, 2005). It is thought that the firing of these cells allow the animal to track its location, and maintain robust cue invariant firing of place cells. Furthermore, it has been shown that the firing pattern of place cells becomes tuned to familiar environments with repeated exposure, and remains consistent over time, responding similarly to different environments with comparable spatial configurations (Lever, Wills, Cacucci, Burgess, & O'Keefe, 2002). This suggests that place cells in the hippocampus encode a long-term memory for particular environments and the locations within them. Whilst place cells have principally been studied in the rat, evidence of these cells has also been reported in the human. Ekstrom et al., (2003) recorded from the medial temporal lobes in epileptic patients undergoing invasive monitoring with intracranial electrodes whilst they navigated virtual reality environments, and found cells in the hippocampus responding to specific spatial locations, and cells in the parahippocampal region responding to views of navigationally relevant landmarks.

The ventral visual pathway is generally thought to allow for invariant recognition, however, the dominant view is that scene processing at the level of the extrastriate visual system is view dependent (Epstein, 2005). The ventral visual pathway however, terminates in the medial temporal lobes and the hippocampus, which are known to contain view invariant representations of the environment (Muller & Kubie, 1987; O'Keefe & Conway, 1978). The functional role of these regions in spatial perception and memory suggest that the ventral stream is involved in establishing view invariant representations of the environment. Information about the geometry of scenes is needed for spatial navigation, and to encode memory for places (Milner & Goodale, 1995). It is necessary for the ventral stream to build a representation that can allow identification of a given scene from any view point, including views that have not been encountered before. Such a representation may be achieved in a single transformation, or in multiple incremental steps. If place selective regions within the ventral visual pathway are encoding information about the geometry of a scene, then they must necessarily establish a degree of invariance to changes in view, and as such, the relative importance of the viewpoint dimension would be expected to decrease as the representation becomes more invariant. The present experiment adopts a parametric adaptation paradigm to investigate the sensitivity in place selective cortical regions to the viewpoint dimension of the representation space of scenes.

### *2.1.1 Visual scene processing in the brain*

The main focus of the present thesis is the ventral surface of the extrastriate visual system; however, the present chapter investigates how the ventral stream establishes view invariant representations of scenes, which necessarily involves spatial processing. It is also necessary to consider other scene selective regions not typically considered part of the ventral stream, such as the retrosplenial complex and the hippocampus, and these regions along with the PPA form a network

involved in spatial navigation, scene perception, and topographical and episodic memory (Byrne & Becker, 2007).

There are two cortical areas that have been shown to be involved in the processing of place stimuli. The first is the parahippocampal cortex within the VOT, and in particular, a functionally defined region within the parahippocampal cortex responding selectively to place stimuli compared to other types of visual stimuli including faces, single objects, and arrays of other objects, the PPA (Epstein & Kanwisher, 1998). The second is the retrosplenial cortex, located towards the anterior end of the calcarine sulcus. There is some ambiguity associated with the definition of this region; functional neuroimaging studies often localise the region by selecting voxels in medial parietal areas showing selectivity for place stimuli. This approach is analogous to the method of functionally localising the PPA, however, the resulting region may differ from the anatomically defined retrosplenial cortex, which includes Brodmann areas 29 and 30. Parts of the posterior cingulate region (BA23) can often be included when using functional localisation. Vann et al. (2009) emphasise the importance of separating the retrosplenial cortex from the posterior cingulate due to cytoarchitectural differences between these regions. When discussing the neuroimaging literature, and the present findings, the terminology used by Epstein et al. (2008) is used, whereby the functionally defined region is referred to as the retrosplenial complex (RSC).

The functional significance and the nature of the representation of visual scenes in these regions have been extensively investigated in recent years. The emerging picture is that the PPA and the RSC play important but distinct roles in both spatial navigation and perception of the local visual environment.

### *2.1.2 The functional role of the PPA*

The PPA has been shown to play a critical role in the perception of scenes. Damage to this region tends to result in general deficits in scene and landmark recognition (Landis, Cummings, Benson, & Palmer, 1986). Furthermore, other

neuropsychological studies have associated damage to this region with deficits in spatial navigation and topographical learning (Aguirre & D'Esposito, 1999; Habib & Sirigu, 1987).

Epstein and Kanwisher (1998) found that the response was equivalent for scenes containing many discrete objects and scenes containing no objects. The response for both of these conditions was greater than for an array of the same objects displayed on a white background, thus the response in the PPA is driven by the presence of local background geometry. The response in the PPA region could not be accounted for by low level image differences between scene stimuli and other stimulus categories. When the scene images were scrambled, preserving the low level image characteristics, but removing high level characteristics indicative of scene stimuli, the effect was extinguished. Furthermore, it was found that fracturing the scene images so that they no longer formed a coherent 3-dimensional geometry also led to a reduction in response. The response in the PPA has been shown to be minimally effected by familiarity with the viewed scene, and as such, the PPA is thought to represent the immediate visual scene, irrespective of any high level semantic associations (Epstein, Harris, Stanley, & Kanwisher, 1999).

The response in the PPA to scene stimuli seems to be largely automatic, and not affected by familiarity with the depicted environment, thus a primarily bottom up process is sufficient to account for the response in this region. The PPA is likely to be involved in the process of building a representation of the environment in order to facilitate spatial navigation and the view invariant representation present in the hippocampus. To understand the functional role of regions involved in spatial cognition, it is necessary to investigate the nature of the representation, or coding in these regions selective for scene stimuli. A key question is the degree to which the PPA can show invariance to changes in the image, in particular, changes in viewpoint.

Several studies have used fMR-a to address the nature of representation in the PPA, and this has led to the dominant view that the PPA processes the scene in a view dependent way (Epstein, 2005). Epstein and colleagues (2003) found that the

PPA showed adaptation to sequentially presented views of the same scene, from the same viewpoint. When different scenes were presented however, no adaptation was found. Crucially, when the same scene was shown rotated in view, the PPA showed no adaptation, and it was concluded that the PPA responded to the new view of the scene in an equivalent manner to its response to a completely novel scene. This finding was later replicated by Epstein et al. (2007) who also examined the adaptation response in the PPA for effects of familiarity with the scenes. It was found that the representation was still view dependent even for highly familiar locations.

The evidence for view dependence in the PPA is not conclusive, as there is evidence that under some circumstances the PPA region can show view invariant adaptation. Using a block design fMR-adaptation experiment, Ewbank et al. (2005) found that the PPA did show significant adaptation that was stable across sequential changes in view point within a block. Ewbank et al. used a change in view of 15° between images within a block design experiment, whereas previous studies used view changes as high as 60° and event related designs. It is possible that these different findings could be accounted for by the amount of view shift used.

### *2.1.3 The functional role of the RSC*

The role of the RSC in scene processing is not yet clear, and considerable recent research attention has been paid to investigating the nature of the representation in this region. Epstein et al. (2007), examined the RSC, along with the PPA, for invariance to changes in viewpoint. It was found that the RSC's pattern of response across conditions was not appreciably different from that of the PPA. They concluded that representation in the RSC is view-dependent. Park and Chun (2009) attempted to address this question by presenting participants with snapshot images from panoramic scene sequences in an event related fMRI design. Attenuation of response to the same depicted location was taken to reflect adaptation due to repetition of location. When the same image was repeated, significant adaptation



was found in both the PPA and the RSC. During panoramic presentation however, only the RSC showed significant adaptation. This was interpreted as the RSC responding in a view invariant manner. A key difference that may account for these conflicting results is that Epstein et al. used view point changes resulting from translation through the environment, whereas Park and Chun used panoramic view point changes showing different scenes around a fixed location. Location within the environment does not change during panoramic view point changes and as such, the RSC may treat each successive image the same. The RSC may be sensitive to perceived location in space, and if so then this manipulation is critical to its response.

A study which further examined the functional role of the RSC and PPA in processing location information was carried out by Epstein et al. (2007). In this study, student participants were asked to make judgements about the location of the depicted scene within their own campus (location), whether the view in the scene faced east or west (orientation), or whether the scene was familiar/unfamiliar. As a control they were also asked to make analogous judgments about object images. It was found the PPA responded equally strongly to all conditions depicting places, whereas the RSC was strongly engaged during the retrieval of location information (both location and orientation). The RSC was also more strongly activated by familiar scenes than unfamiliar. These results support the idea that the PPA is involved in the perception of scenes, regardless of task, whereas the RSC is involved in the retrieval of spatial information from memory.

Studies up to now have shown that the PPA and the RSC play distinct roles in the processing of visual scenes. Converging evidence from the neuropsychological and neuroimaging literature suggests that the principal role of the PPA is primarily perceptual. The PPA shows considerable activation when perceiving any scene like stimuli, and this activation is largely unaffected by the kind of task being undertaken, or by task difficulty. The RSC however, appears to play a distinct role in the processing of scene stimuli compared to the PPA and may represent more abstract information, not dependent on the visual properties of the scene.

### 2.1.4 *Summary and aims*

The focus of the present chapter is the representation of the viewpoint of landscapes within place selective cortex. The aims of the chapter are three-fold. The first is to show that invariance is not an all or none phenomenon, and that fMR-adaptation can be used as a measure of invariance within category selective cortex. The second is to go beyond what has been done before in describing the extent to which place selective brain regions are invariant, or sensitive, to changes in view. A final aim is to investigate the functional significance of place selective brain regions by testing for links between view invariance in these regions and topographical memory for landscape scenes.

Investigating the nature of the representation is an important step in understanding the function of scene selective brain regions such as the PPA and RSC. Previous experiments have manipulated view shift and compared the amount of adaptation in place selective cortex when there is some change in view point, to a condition where there is no change. A key difference between studies that find view invariant representations of scenes, and those that do not, is that the former used smaller view shift manipulations. Such studies rely on an assumption that a given representation will be either view invariant, or view dependent. If view invariance is established in a gradual manner, then the degree of viewpoint change between successive images in a block could play a role in determining the degree of adaptation.

In the present study a block design similar to that used by Ewbank et al.(2005) was carried out with the addition of a parametric manipulation of view shift. Four view shift conditions are used ( $0^\circ$ ,  $5^\circ$ ,  $10^\circ$ ,  $15^\circ$ ). All of these conditions are either the same, or smaller than, the largest shift condition used by Ewbank et al., and in this study, view invariant adaptation was reported. If place selective regions such as the PPA and RSC are invariant to view shifts of this size, and adaptation is an all or none phenomenon, then equivalent amounts of adaptation will be seen for all conditions. If place selective regions of cortex show an increase

in response with increasing view shift then these brain regions are sensitive to the *degree* of view shift and a parametric adaptation paradigm can be used to measure sensitivity to the high level property of view shift.

A parametric manipulation of the view shift of landscape stimulus can allow the hypothetical representation space of scenes to be sampled along a key dimension, viewpoint. Participants viewed blocks of images depicting landscape scenes, the degree of view change between successive images within a block increased parametrically across conditions. When the degree of view shift within a condition is small, then the sampling distance along the view shift dimension is small. When the degree of view shift is large, distance along the view shift dimension is large. The degree of adaptation found within place selective cortex for a given level of view shift, reflects the sensitivity of the neural representation to the dimension of viewpoint. Different views of the same place are hypothesised to activate similar representations in place selective cortex, resulting in fMR-adaptation relative to different scenes. However, the degree of adaptation was predicted to vary depending upon the size of view shift.

The final aim relates to the functional significance of place selective brain regions. It is necessary given the nature of the present adaptation design, that the task be limited to passive viewing of scene stimuli in the scanner. We can however, examine the role that scene selective regions play in more complex cognitive processes, such as topographical memory, by investigating links between neural response and behavioural performance outside the scanner. The stimuli used in the fMRI experiment were designed to be equivalent to those of Hartley et al. (2007), who developed a test of topographical short term memory to investigate the involvement of the hippocampus in topographical memory. They compared patients with hippocampal damage and normal controls on this test and found that hippocampal damage resulted in impaired topographical memory. In order to perform well on the task, it is necessary to form a rich perceptual representation of the landscape stimuli, as well as to compute some form of view invariant representation and store it in memory. This same test of topographical memory was

administered to all the current participants before scanning. The PPA and the RSC have been suggested as potential sites of these perceptual and memory encoding processes, and as such it was predicted that response in these regions would show a positive relationship with topographical memory behavioural performance. Furthermore, the degree of view sensitivity in these regions was predicted to be related to topographical memory performance.

## **2.2 Methods**

### *2.2.1 Participants*

21 healthy right handed participants took part in the study (12 male and 9 female) with a mean age of  $25 \pm 3.6$ ; all had normal, or corrected to normal vision, they were drawn mostly from the University of York postgraduate population. Written consent was obtained from all participants and the study was approved by the York Neuroimaging Centre Ethics Committee.

### *2.2.2 Imaging parameters*

All experiments were carried out using a General Electric 3 Tesla HD Excite MRI scanner at the York Neuroimaging Centre (YNiC) at the University of York. An 8 channel, phased-array head coil (GE, Milwaukee) tuned to 127.4 MHz was used to acquire MRI data. Functional images were obtained using a gradient-echo EPI sequence (TR 3000ms, TE 25.6ms, flip angle  $90^\circ$ , matrix 128 x 128, field of view 28.8cm) with 38 contiguous axial slices at a resolution of 2.25x2.25x3mm. Functional images were centred over the occipital cortex in the plane of the temporal lobes. Whole brain T1 weighted structural scans were obtained at 1mm isotropic resolution (TR 7800ms, TE 2900ms, flip angle  $20^\circ$ , field of view 29cm). A T1-weighted flair image was also taken in the same plane as the EPI sequence to aid registration.

### *2.2.3 Landscape stimuli*

The landscape stimuli were based on a heightfield (a 2D array of altitudes) constructed using MATLAB 6.1 (Mathworks Inc.). Each stimulus heightfield was built by summing six 256x256 height fields. There were four hills of varying shapes and sizes, placed at different locations around the centre, and a semicircular mountain range in the background of the stimulus images. Stimuli consisted of 3d rendered place stimuli and a smoothed interpolated 2D noise field. The noise served to add

realism and additional unique small scale topographical features. For further details of the stimulus creation and image rendering see Hartley et al. (2007).

A modified version of the landscape stimuli from the behavioural task were created for use in the scanner. The spatial configuration of each scene was retained but the texture, lighting and colour properties were made uniform for all images in all conditions. For each scene, 27 unique views were created; each view was separated by 5°. The light source was a fixed 110° from the main viewing angle in each image in order to insure that places were equivalent in terms of average luminance throughout each block.

### 2.2.4 *Behavioural task*

The task used was the same as the topographical memory task used by Hartley et al. (2007); for full details of the procedure see this paper. The stimuli were presented using E-Prime (<http://www.pstnet.com/products/e-prime/>) and participants carried out a 30-trial, delayed match to sample task. Participants were shown a single landscape image for eight seconds, followed by a blank screen for two seconds; they were then shown a display with four landscape images and asked to choose the previously viewed place. The target depicted the same geographical layout but varied in terms of viewpoint, texture and lighting properties (designed to reflect time of day/year changes in the landscape). The three non-target images also varied in terms of the spatial configuration of the landscape depicted. Participants were given a maximum of 20 seconds to respond.

### 2.2.5 *Localiser scan*

Subjects viewed 20 blocks of 10 images. Each block contained images from one of five different categories: faces, bodies, objects, places, or Fourier scrambled images of the former categories. Face images were taken from the Psychological Image Collection at Stirling (PICS; <http://www.pics.psych.stir.ac.uk>) and body images were taken from a body image collection at Bangor (Downing et al. 2001; <http://www.bangor.ac.uk/~pss811/page7/page7.html>). Images of other categories

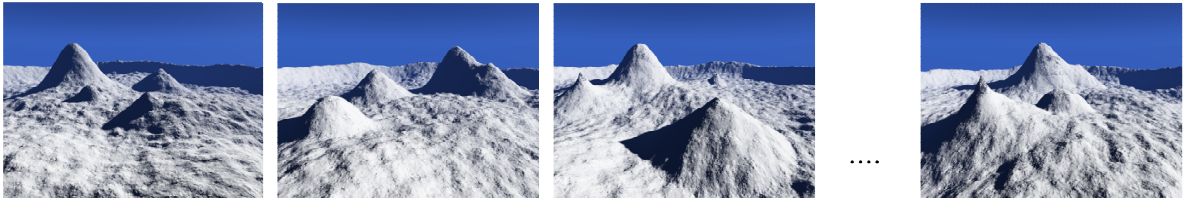
were taken from a variety of web-based sources. Each image was presented for 700ms followed by a 200ms fixation cross. Stimulus blocks were separated by a 9s fixation screen. Each condition was repeated four times, and arranged in a counterbalanced block design.

### 2.2.6 *Adaptation scan*

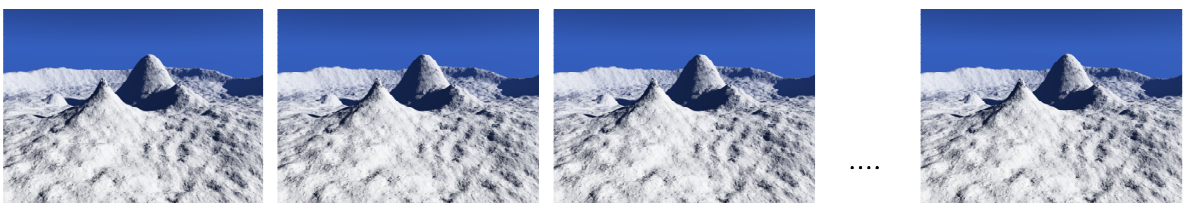
During the adaptation scan, participants viewed blocks of images of computer generated landscapes. In different experimental conditions the blocks could be comprised of identical views of the same place, different views of the same place, or different places (i.e., landscapes with distinct topographical features). For blocks of the images showing the same place, the degree of view shift between successive images was manipulated to create six experimental conditions in total, with eight blocks per condition. 30 unique places were used, each place was used at most twice and no unique image from any of the places appeared more than once (with the exception of the same image condition). The six conditions were: (1) Different places, random views selected from a pool of ten places, images from the same place never occurring more than once within each block; (2) Same image, same place, 0° view shift; (3) Same place, 5° view-shift, sorted; (4) Same place, 10° view-shift, sorted; (5) Same place, 15° view-shift, sorted; (6) Same place, shuffled sequence with a 10° average view shift (**Figure 2.1**). The shuffled sequence condition was designed as a control for the predictability of the image sequence, which was high in all sorted view shift conditions. Unpredictable image sequences were generated with an average view shift throughout the block of 10°. This was achieved by generating unique random permutations of view shift orders and selecting those with appropriate mean differences between successive shift sizes. An entirely unique shift-sequence was used for each block in the shuffled condition. For all view shifting conditions (2-6), each landscape scene was used twice within each condition, where possible, the range of images sampled from a given landscape were chosen to be from a range of previously unused views. There were 48 blocks in total, which were presented in a pseudo-random order throughout the

functional run. Each image was presented for 800ms and separated by 200ms fixation cross. Each block consisted of 9 images, lasting 9s in total. The blocks were separated by fixation cross interval lasting 9s.

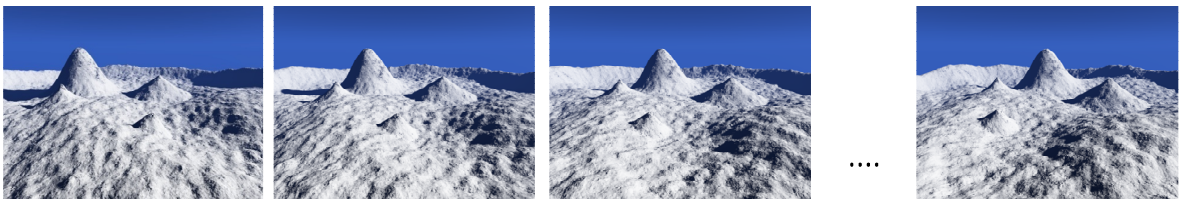
**Different landscapes:** random views



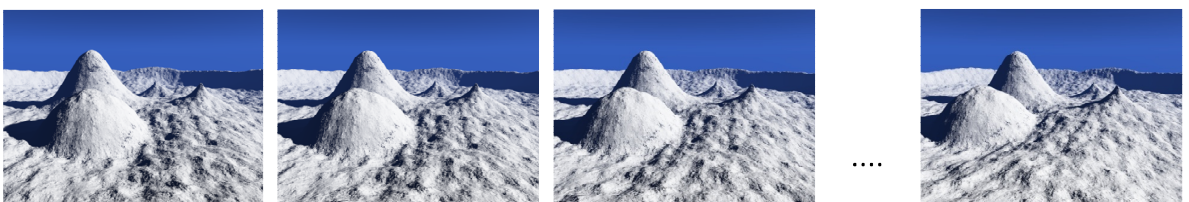
Same image: same landscape,  $0^\circ$  view shift



Same landscape,  $10^\circ$  view-shift, ordered sequence



Same landscape,  $10^\circ$  average view-shift, shuffled sequence



**Figure 2.1.** Examples of conditions in the adaptation scan. *Different landscapes (top row), each image within the block depicts a novel landscape from a randomly chosen view. Same landscape, same image (second row), a single view of a landscape repeated throughout the block. Same landscape,  $10^\circ$  view shift (third row), each image progresses around the landscape in an ordered sequence of  $10^\circ$  steps. Same landscape,  $10^\circ$  shuffled (bottom row), sequence moves back and forth around the landscape unpredictably yet maintain an average shift throughout the block of  $10^\circ$ . Also included were  $5^\circ$  and  $15^\circ$  view shift blocks (not shown).*



### 2.2.7 *fMRI attention task*

In order to maintain attention throughout the experiment, participants were asked to respond to a red dot superimposed at a random location on the image. Either one or two dots were present in each block and the distribution of one dot and two dot blocks was uniform across conditions. The red dot detection task was used in both the localiser and experimental scan. Results of the task were used primarily to verify attention and were not considered in the main analysis. An accuracy level lower than 80% on the red dot task was deemed to be an appropriate exclusion criteria, no participant's score fell below this threshold.

### 2.2.8 *fMRI analysis*

#### *Whole brain*

Whole brain analysis of the fMRI data was carried out using FEAT (<http://www.fmrib.ox.ac.uk/fsl>; Smith et al., 2004). The initial 9 s of data from each scan were removed to minimize the effects of magnetic saturation. Motion correction was performed, along with removal of non-brain structures, spatial smoothing (Gaussian, FWHM 6mm) and temporal high-pass filtering (cut off, 0.01 Hz). A two-step mixed-effects analysis was carried out. An initial fixed effects model of each of the six conditions relative to inter-block fixation baseline was used to analyse individual participant data. Group analysis was carried out using a random-effects model (FLAME, <http://www.fmrib.ox.ac.uk/fsl>). Registration to high-resolution individual space and standard MNI space was carried out using FLIRT (FMRIB's Linear Image Registration Tool) for each individual (Jenkinson, Bannister, Brady, & Smith, 2002). Contrasts were calculated for each view shift condition relative to a baseline of the different place condition; significant activity in these contrasts is considered to reflect adaptation due to the same depicted landscape scene. A contrast of the 10° view shift condition with the shuffled sequence condition was calculated in order to show differences in response due to shift-sequence predictability. A contrast of all landscape blocks vs. fixation baseline was

calculated. In order to detect voxels which show a view sensitive pattern of response, a linear trend contrast was computed across all four view shift conditions (same image, same place condition was considered as a 0° view-shift). The effect size within a voxel for this contrast reflects the extent to which activity increases linearly with increasing view shift. Such a contrast does not show activity relative to a baseline condition and as such is sensitive to trends across conditions occurring at all absolute activity levels. Unless otherwise stated, whole brain images are presented with a threshold of  $P < 0.001$  (uncorrected).

### *Behavioural regressor*

Parameter estimates derived from the whole brain analyse were entered into a separate model at the group level, in which individuals' scores on the topographical memory test were entered into a simple regression model. Individual topographical memory scores were transformed into a deviation from overall mean score ( $[x^i - \text{mean}(x)] / \max(x)$ ). Two contrasts were examined for correlations with behavioural performance; activity in the linear trend contrast, considered to reflect a measure of view sensitivity at the voxel level, and a contrast of all six conditions in the adaptation scan relative to fixation baseline (active > rest), considered to reflect landscape processing generally.

### *Regions of interest*

The independent functional localiser scan was used to identify regions of interest including place selective cortical regions (PPA and RSC) and object selective LOC (considered as a category-selective non-spatial control region and for later comparison with other stimulus types). As a low-level visual control region, primary visual cortex (V1) was identified. The contrast of places > faces from the localiser scan was used to define the PPA and RSC ROI. Objects > scrambled was used to define the object selective LOC. All functionally localised regions were defined in each individual by anatomically constraining the activation using probability maps from the Harvard-Oxford cortical structural atlas

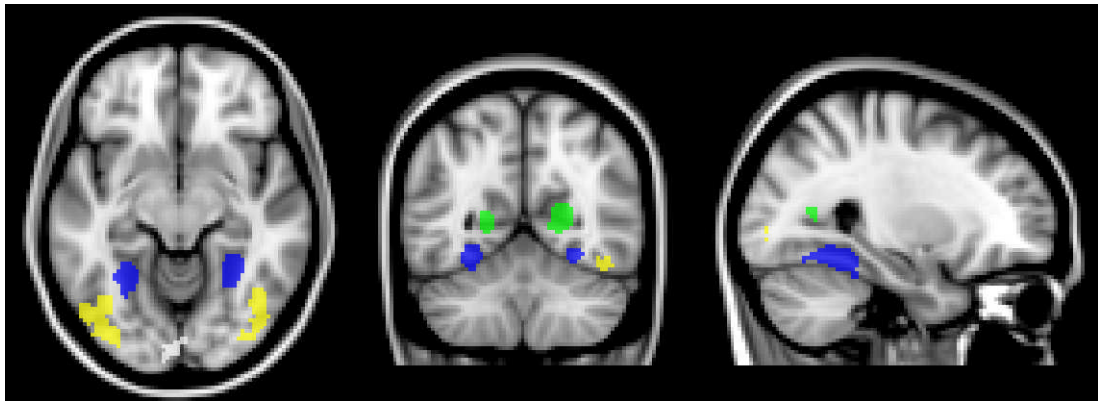
(<http://www.fmrib.ox.ac.uk/fsl/fslview/atlas-descriptions.html>), and selecting up to the 50 most active voxels within this mask. A minimum threshold of  $P < 0.001$  (uncorrected) was applied, and individuals not showing activation above this threshold for a given region were excluded from the analysis. Three individuals failed to show a left RSC, all other regions were localised bilaterally in all participants. The V1 ROI was defined anatomically by masking the calcarine sulcus of each individual structural scan, and further constraining the mask using active > rest in the localiser scan, ensuring that selected voxels were responsive to the area of the visual field stimulated throughout the experiment.

GLM parameter estimates for each adaptation scan condition were extracted for each ROI and all voxels within each region were averaged together. Activation values for each region were converted from arbitrary units to units of % signal change relative to fixation baseline ( $\text{parameter estimate} / 100 * \text{Regressor Height}$ ). Responses from all ROI's were subjected to repeated measures ANOVA to determine significant differences between stimulus conditions, regions and hemispheres.

## **2.3 Results**

### *2.3.1 Localiser*

Place and face selective regions of interest were identified within the (VOT) region of the brain (**Figure 2.2; Table 2.1**). The PPA was identified as a region along the parahippocampal gyrus that showed a greater selectivity for place stimuli relative to face stimuli (Epstein & Kanwisher, 1998). The Retrosplenial Complex (RSC) was defined using the same Places>Faces contrast but the activation was constrained anatomically to the Retrosplenial area (see methods). The RSC defined in this manner may include parts of the posterior cingulate, and as such should be distinguished from the anatomical retrosplenial cortex (see Epstein, Parker, & Feiler, 2007 for discussion). The LOC was identified as a region in the lateral occipital cortex or posterior cingulate showing a greater selectivity for objects relative to scrambled objects (Malach et al., 1995). Both PPA and LOC contrasts elicited sizable bilateral regions of activation in all participants, whereas the RSC showed more variation in size and was not identified at all in the left hemisphere of three participants. The V1 ROI was created anatomically by manually identifying the calcarine sulcus in each individual and constraining the region using a functional contrast of active>rest in the localiser scan. This ensured that only active portions of the visual field were sampled. Masks were created for each of the above regions which were then used to extract mean time series' for each region within the adaptation scan for all six conditions.



**Figure 2.2.** Category selective, localiser scan ROI masks shown in MNI space. Blue=PPA (places>faces), green=RSC (places>faces), yellow=LOC (objects>scrambled), white=V1 (anatomical calcarine sulcus combine with active>rest). Masks are created using a minimum threshold of  $P < 0.001$  (uncorrected). Images are centred on MNI coordinates: 28, -58, -10. Scans are displayed using neurological convention with the right hemisphere displayed on the right.

**Table 2.1.** MNI coordinates of regions of interest peaks

Area	Hemisphere	MNI Co-ordinates			Z-value
		X	Y	Z	
PPA	Right	30	-42	-16	7.2
	Left	-28	-50	-14	7.85
RSC	Right	18	-54	2	6.99
	Left	-16	-62	2	5.34
LOC	Right	46	-66	-18	6.09
	Left	-50	-72	-14	6.78
V1	Right	16	-92	6	5.67
	Left	-16	-92	6	5.1

### 2.3.2 Predictability

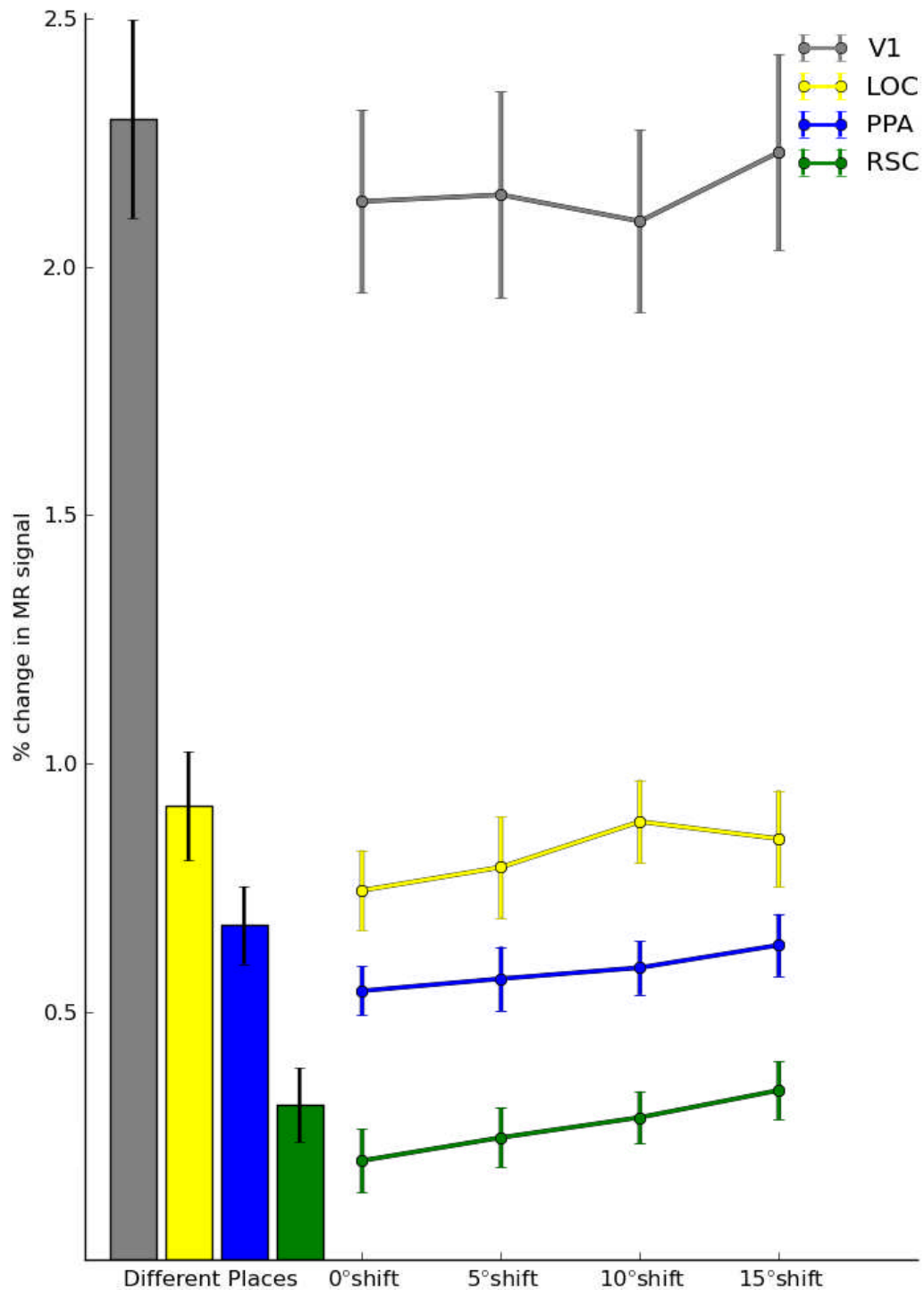
It is possible that differences between the different place condition and any view shifting condition simply reflect differences in the predictability of the shift sequence. In order to control for this potential confound a condition was included whereby the shift sequence was unpredictable. The size of the view shift was varied randomly but the average view shift was maintained at  $10^\circ$ , thus the condition was

treated as directly comparable to the 10° view shift condition. Planned comparisons showed no difference between the 10° view shift sorted and 10° shuffled conditions in place selective ROI's, and no interaction was found between this contrast and either hemisphere or region ( $F_s < 1$ ).

A whole brain comparison of the 10° ordered sequence condition and the 10° shuffled condition was carried out and no significant differences were found. Unpredictable sequences were not associated with different levels of activation anywhere in the brain. The 10° shuffled condition was not considered further.

### *2.3.3 Adaptation*

No main effects or interactions involving hemisphere were found in any ROI, responses across hemispheres were pooled for all subsequent analysis. The responses for all of the view shift conditions relative to the different condition can be seen in **Figure 2.3**. Planned adaptation comparisons for each view shift condition compared to the different places condition was carried out for all regions (**Table 2.3**).



**Figure 2.3.** Response to all view shift conditions in all ROIs.

Bars represent the response to different place condition; connected line represents the response across view shift conditions. Data are shown for PPA (blue), RSC (green), LOC (yellow), and V1 (grey). Error bars represent  $\pm 1$  standard error of the mean. All category selective regions show a significant linear increase in response to increasing view shifts whereas V1 does not.

**Table 2.2.** *Different-Same adaptation contrasts, all ROIs.*

ROI	Adaptation Contrast:- Different - Same (0°,5°,10°,15°)
PPA	$F_{1,20}=5.988, P<0.05$
RSC	$F_{1,20}=1.157, P=0.295$
LOC	$F_{1,20}=3.889, P=0.063$
V1	$F_{1,20}=9.861, P<0.01$

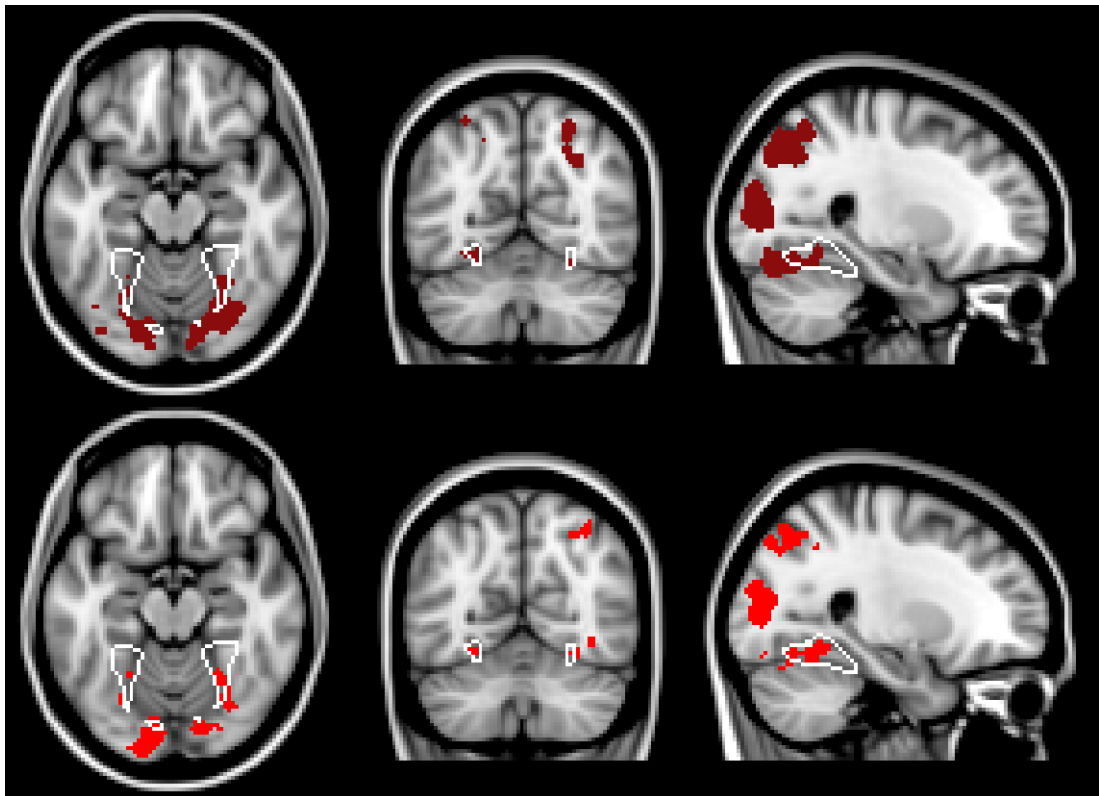
**Table 2.3.** *Adaptation contrasts for each view shift condition and all ROIs.*

ROI	Adaptation Contrast	
PPA	Different - 0°	$F_{1,20}=8.037, P<0.01$
	Different - 5°	$F_{1,20}=6.177, P<0.05$
	Different - 10°	$F_{1,20}=3.216, P=0.08$
	Different - 15°	$F_{1,20}=1.439, P=0.24$
RSC	Different - 0°	$F_{1,20}=6.683, P<0.05$
	Different - 5°	$F_{1,20}=2.126, P=0.16$
	Different - 10°	$F_{1,20}<1, P=0.664$
	Different - 15°	$F_{1,20}<1, P=0.455$
LOC	Different - 0°	$F_{1,20}=7.723, P<0.05$
	Different - 5°	$F_{1,20}=3.610, P=0.08$
	Different - 10°	$F_{1,20}<1, P=0.634$
	Different - 15°	$F_{1,20}=3.230, P=0.09$
V1	Different - 0°	$F_{1,20}=5.159, P<0.05$
	Different - 5°	$F_{1,20}=5.854, P<0.05$
	Different - 10°	$F_{1,20}=7.531, P<0.05$
	Different - 15°	$F_{1,20}=4.097, P=0.057$

Same identity adaptation was found in both V1 and the PPA (**Table 2.2**). **Table 2.3** shows the adaptation contrasts for each view shift condition and all ROIs. In V1, adaptation was shown for all view shift conditions, however the magnitude of the adaptation is roughly the same regardless of the size of the view shift. All three extrastriate, category selective regions show greater adaptation for lower view shifts and a release from adaptation at higher ones.



Whole brain adaptation responses for both static and view shifting landscape scenes are shown in **Figure 2.4**. Regions shown in dark red respond greater to the different place condition relative to the same place, static view shift condition. Regions in light red which respond more to the different place condition than to shifting views of the same place. For this comparison all three view shift conditions were combined together. The view of the depicted scene was changing within the block and as such adaptation shown is unlikely to reflect image level adaptation. The outline of the PPA from the localiser scan is overlaid on the figure and it can be seen that the adaptation to the static scene condition is more extensive immediately posterior to the PPA.



**Figure 2.4.** Same place adaptation.

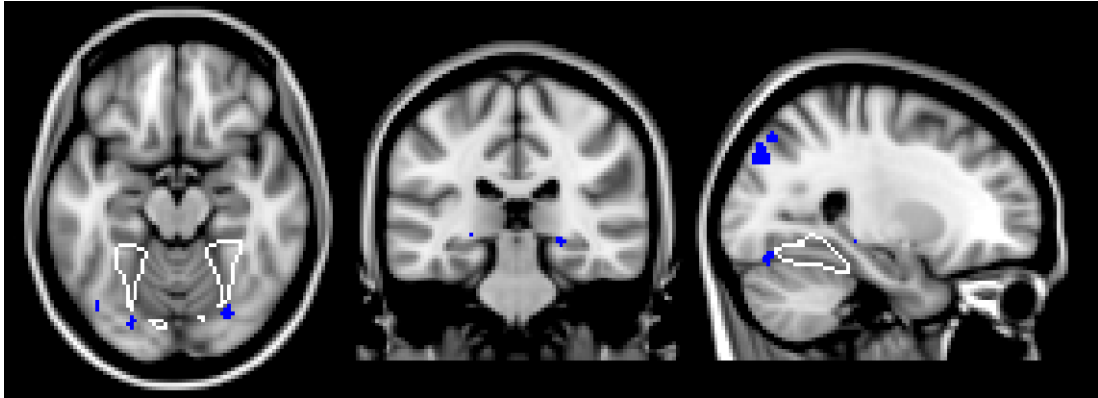
*(Upper) Areas in dark red are voxels showing adaptation to repeated presentation of the same place as defined by the contrast different places > same place (0° view shift). (Lower) Areas in bright red are voxels showing adaptation to the same place regardless of view shift, as defined by the contrast different places > shifting place (5°, 10°, & 15° view shifts). Outlines of the PPA as defined by a contrast of Places > Faces are shown in white. Statistical images are created using a threshold of  $P < 0.001$  (uncorrected). Images centred on MNI coordinates: 26, -62, -16. Scans are displayed using neurological convention with the right hemisphere on the right.*

### 2.3.4 View shift sensitivity

Scenes depicted in all view shift conditions maintained the same identity throughout the block, yet these conditions differed only in the degree of view shift steps throughout the block. It was thought that a linear increase in view shift would result in a similar linear increase in BOLD response in place selective regions, but not in non-place selective regions. In order to look for differences in view shift response across all ROIs, a 4x4 repeated measures ANOVA was carried out (region, view shift). There was a strong main effect of region ( $F^{(3,60)} = 69.465$ ,  $P < 0.001$ ), but no overall main effect of view shift ( $F^{(3,60)} = 2.332$ ,  $P = 0.083$ ). There was however, an interaction of region with view shift ( $F^{(9,180)} = 2.154$ ,  $P < 0.05$ ). See the line graph section of **Figure 2.3** for an illustration of these effects.

The same image condition was treated as a 0° view shift and a linear trend contrast across the four view shift conditions was computed. This was examined independently for each region. Responses in the place selective PPA and RSC regions were found to be modulated by degree of view shift. A significant linear trend was found in the PPA ( $F^{(1,20)} = 7.024$ ,  $P < 0.05$ ) and the RSC ( $F^{(1,20)} = 12.251$ ,  $P < 0.005$ ). Significant linear trends were also found in the non-place selective LOC ( $F^{(1,20)} = 8.702$ ,  $P < 0.01$ ). However, no linear trend was observed in the V1 ROI ( $F^{(1,20)} < 1$ ,  $P = 0.507$ ).

A whole brain linear trend contrast was carried out in order to identify regions that show a linear increase in activation with parametric increases in view shift. Significant activation in the linear trend contrast was found bi-laterally in the hippocampal formation as well as regions immediately posterior to the boundaries of the PPA (**Figure 2.5**). Additional activation in this contrast was found bilaterally in superior lateral occipital cortex approximately corresponds to reported coordinates of area V5/MT (Dumoulin et al., 2000).



**Figure 2.5.** *Linear responses to view shift.*

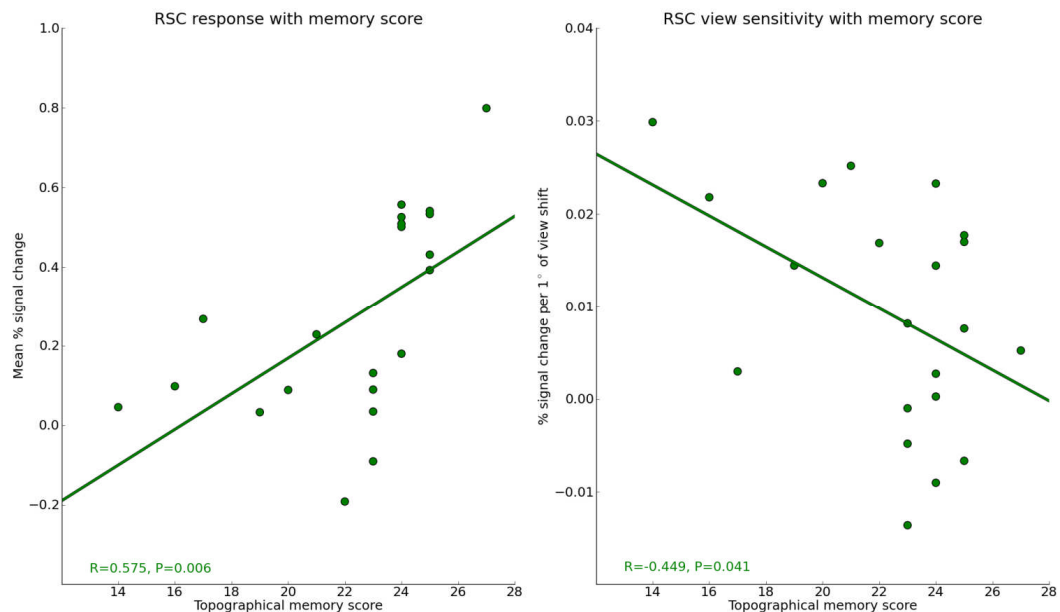
*Bilateral activation in the hippocampus that shows a linear increase in activation with successively increasing view shifts (blue). This contrast was carried out using the 0°, 5°, 10°, & 15° conditions with a contrast vector -3, -1, 1, 3. Statistical images are created using a threshold of  $P < 0.001$  (uncorrected). Images are centred on MNI coordinates corresponding to the medial portion of the hippocampus: 26, -30, -16. Scans are displayed using neurological convention with the right hemisphere displayed on the right.*

### 2.3.5 Behavioural measures

All participants performed to an acceptable standard in the red dot task for both the adaptation scan and localiser scan. The mean score in the adaptation scan was  $95.62 \pm 6.43$ , in the localiser the mean score was  $96.28 \pm 5.79$ . One participant's localiser red dot scores were lost due to a technical fault; the subject reported no problems performing the task and performed above average for the adaptation scan. All participants were considered to be paying sufficient attention to the stimuli. The average score in the topographical memory task (out of a maximum of 30) was  $22 \pm 3.3$  with all participants performing well above chance. The minimum score was 14 and the maximum was 27.

All ROIs were examined for correlations between BOLD activity (active > rest) and topographical memory performance. Only the RSC showed a significant positive relationship with performance, which was found to be highly significant in ( $r = 0.575$ ,  $P < 0.005$ ,  $N = 21$ ). Further correlations were found between RSC response and view sensitivity, which was quantified using the slope value of a least squares linear fit across all four view shift conditions for each individual. There was a negative relationship between view sensitivity and behavioural performance ( $r = -$

0.449,  $P < 0.05$ ,  $N = 21$ ). View sensitivity showed no significant relationship with behavioural performance in any other region. These results indicate that individuals with better memory performance have more active RSC regions, and that greater view sensitivity in RSC is predictive of poorer topographical memory performance. See **Figure 2.6** for scatter plots.



**Figure 2.6.** Scatterplots of brain response and behaviour.

(left) The relationship in the RSC between topographical memory performance and BOLD response in passive viewing of place stimuli. (right) The relationship in the RSC between topographical memory performance and slope of a least squares fit line across four view shift conditions.

In addition to looking for behavioural correlations in all functionally defined ROIs, scores were transformed into a deviation from the mean score and entered as a behavioural regressor into the whole brain analysis. Parameter estimates for each of the six conditions and the whole brain linear trend contrast (reflecting view sensitivity) were examined for correlations with topographical memory. Marginally significant correlations with memory scores were observed bilaterally at just below the critical threshold in the medial portion of the hippocampus, and within or nearby the RSC. An additional marginal sub-threshold negative relationship was observed bilaterally between view sensitivity and topographical memory

performance in voxels roughly corresponding to the RSC. Due to the marginal significance of these effects the whole brain maps are not shown, however, it is thought that with greater statistical power the pattern of brain areas showing relationships with behaviour would correspond well to the patterns observed in the ROI analysis. The additional weak behavioural correlations in the hippocampus could not be examined using ROI methods; however, they may be informative when considering the role of the hippocampus in topographical memory and spatial navigation.

## **2.4 Discussion**

### *2.4.1 Representation of scenes in place selective cortex*

The primary aims of the present chapter relate to the nature of the representation of landscape scenes in place selective brain regions. Some previous studies have reported adaptation in the PPA that is robust across changes in view (Ewbank et al., 2005), whereas others reported that changes in view resulted in a release from adaptation (Epstein et al., 2003). The present study tested whether the amount of view change in the images could account for these different previous findings. A parametric manipulation of view shift was carried out in which participants encountered view shifts ranging from 0° up to 15°. It was found that the response in PPA and the RSC regions were sensitive to view shift, showing more adaptation in the lower view shift conditions than the higher. The brain response in these regions was found to increase linearly as view shift increases.

Previously, it has been suggested that the role of the PPA is in the representation of the immediate visual scene. The reason for this is that there is little difference in the adaptation response of the PPA between familiar and unfamiliar environment stimuli (Epstein et al., 2007). The present experiment lends support to this idea, the PPA representation may show view invariance only to small amounts of view change in the scene. This is in line with the amount of change that an individual might encounter as they move around an environment. Therefore the degree of view invariance in the PPA gives rise to a percept of coherence and constancy for the immediate visual scene as one moves through it. The PPA is engaged primarily with the perception of the visual environment, and additional regions must be engaged in order to form more view invariant representations necessary for recognition across larger view shifts that are less likely to occur naturalistically. These might be recruited only when required for the task at hand (e.g., encoding locations and their spatial layout for long-term memory and future navigation).

Other adaptation studies looking at view invariance in the PPA also differed in terms of their methodology. Some used event related designs in which a single image was presented followed by a further image from a different view. Others used block designs similar to the present study, whereby blocks of scene images were presented sequentially, each changing in view by a fixed amount. The sequence in such a block design is entirely predictable, and this predictability may facilitate the visual system identifying each successive image as depicting the same scene. The present study included a condition in which the mean view shift was  $10^\circ$ , however each individual view change varied between  $5^\circ$  and  $15^\circ$  in either direction, and as such was unpredictable. A comparison between this condition and the  $10^\circ$  shift condition revealed no differences anywhere in the brain including the scene selective ROIs. Adaptation was found to the same extent when the sequence is unpredictable, as to when it is predictable. It is therefore suggested that predictability differences between prior studies did not give rise to the differences in adaptation response previously observed.

### *2.4.2 Non-place selective VOT regions*

The pattern of response in the LOC did show that this region is modulated by the degree of view shift for landscapes. The extent of the adaptation was weaker than other regions, and the pattern across same identity view shift conditions does not appear entirely linear. Although this region is not scene selective, it does show some same identity adaptation, and the adaptation shows some sensitivity to view shift. View invariant adaptation for scenes in LOC is not unprecedented (Ewbank et al., 2005). It is perhaps unsurprising that the LOC may also be engaged by scene stimuli as recognition of a scene involves coding the identity of objects and the spatial relationships between them (Milner & Goodale, 1995). It could be suggested that non-linearity in the pattern of response seen across view shift conditions in the LOC arise from different amounts of view change for different objects placed throughout the scene. In the case of the present experiment mountains positioned

close to the viewer make more extreme translations than those in the distance, and this effect would be more pronounced in higher view shift blocks.

If the PPA is involved in representing spatial relationships between objects, it would not be predicted that this region would be sensitive to the viewpoint of objects in isolation. Further investigation could reveal dissociations in view sensitivity responses in VOT regions by examining similar parametric manipulations of view shift in other object categories such as faces or objects. Previous research has shown different degrees of view invariance within the LOC region, with the posterior region LO showing a view dependent pattern of response and the more anterior posterior Fusiform (pF) showing invariant adaptation to changes in view (Grill-Spector et al., 1999; Kourtzi et al., 2003). A parametric adaptation paradigm could be applied to objects in isolation to investigate the degree of sensitivity to view shift for objects in both object selective and non-object selective VOT regions (see chapter 3). Such an approach has the potential to reveal the extent to which view invariance is related to category selectivity.

### *2.4.3 The effects of low level visual change*

A limitation of the present experiment is that low level image features varied across conditions in a similar way to our experimental manipulation of view shift. The low level image properties varied between images in the different condition more than they did for the view shift blocks, and the degree of change in low level properties was greater for high view shift blocks than for low view shift blocks (see **Appendix A. 2**). This is a problem not unique to the present experiment, and applies to most fMR-adaptation experiments investigating invariance. To overcome this we considered the response to our stimuli in V1. It was found that no sensitivity to view shift was observed in primary visual cortex and the region showed a generally higher response to the different condition than all same identity scene conditions. This is informative about the contribution of the low level image properties in driving responses throughout the visual system. If low level changes alone could



account for the differences we observe in VOT regions, then a similar, pattern of view sensitive responses should be observed in V1, which is thought to be largely unresponsive to high level stimulus properties.

Previous studies have shown that scrambled versions of category stimuli, retaining spatial frequency information from their unscrambled counterparts, can illicit some adaptation in VOT regions (Andrews, Clarke, Pell, & Hartley, 2010). The adaptation found to scrambled stimuli was considerably less than that of unscrambled. As such, it has been suggested that the low level image properties can account for some proportion of VOT category selectivity.

Further research is needed to investigate the extent to which low level image features contribute to the effects observed here. Scrambling in ways which vary in similar ways to the present view shift blocks could shed light on this issue. An example of this would involve scrambling a view shift block such that the degree of image change between successive images is equivalent to the change in an intact landscape block. This would make it possible to measure adaptation in high level VOT regions driven by such low level image properties, and to measure adaptation over and above this for intact stimuli. Such a control manipulation is incorporated into the design of chapter 3, when the parametric adaptation paradigm is applied to real world object stimuli.

#### *2.4.4 Links between topographical memory and brain response*

A further aim of the present chapter was to consider the role of place selective regions in topographical memory processes. Activation in the RSC region showed greater sensitivity to view shift than any other region as well as a positive correlation of overall BOLD response with topographical memory performance. View sensitivity of the RSC was also shown to be negatively related to topographical memory performance. The RSC of individuals who performed better in the topographical memory test showed greater absolute response and more view invariance across view shift conditions within this region. This is suggestive of the

RSC's functional significance for recognition of scenes when encountered from different viewpoints and under different viewing conditions.

Activation of the RSC is common in a wide variety of neuroimaging studies. It has been shown, along with other key brain regions, to be tonically active when participants are not performing any task, and therefore has been suggested as a key part of the default mode network (Buckner, Andrews-Hanna, & Schacter, 2008). Two functions with which the RSC is unambiguously involved are episodic memory and spatial navigation (Svoboda, McKinnon, & Levine, 2006; Vann et al., 2009). In the animal literature lesion studies in both the rat and monkey have shown that the retrosplenial cortex is necessary for a wide array of navigation and spatial memory tasks (Aggleton & Vann, 2004). Similar findings about the function of the RSC have emerged from the human neuropsychological literature, where damage to the region has been associated with general episodic memory deficits as well as spatial memory impairments (Rudge & Warrington, 1991). In some cases, specific, selective deficits in spatial navigation have been demonstrated. Maguire (2001), reported patients with retrosplenial lesions that were able to accurately recognise and identify familiar landmarks but are unable to recruit this information to aid navigation. Such patients often also show impairments when asked to learn new routes for the first time.

There is considerable difficulty in establishing the precise function of the RSC region whilst accounting for its role in such a diverse array of cognitive tasks. One attempt to account for the RSCs involvement in both memory and spatial navigation has been put forward by Byrne and Becker (2007). They present a computational model in which the RSC is a central hub with reciprocal connections with low level visual areas and hippocampal and frontal systems. They propose that the role of the RSC is to transform perceptual representations of the environment between egocentric and allocentric coordinate frames in order to facilitate encoding and retrieval in memory. In the present experiment the RSC region is engaged to perform view invariant scene processing and the degree of view sensitivity in this region is predictive of topographical memory ability. Wolbers & Buchel (2005) also

found that the activity in the RSC increased across multiple scanning sessions in which participants formed survey representations of a virtual reality environment, and was predictive of behavioural measures of survey knowledge. The present findings are therefore consistent with an account of the RSC as being involved in computing view invariant representations of the scene in order to subserve recognition from novel view points.

### *2.4.5 View sensitivity outside the VOT*

An exploratory whole brain analysis was carried out to look for areas outside functionally defined regions showing view shift sensitivity. This revealed bilateral regions in the hippocampus, regions posterior to the PPA, and regions in the lateral occipital cortex thought to correspond to area V5/MT which showed a linear increase in response between same scene activation and increasing view shift. In order to interpret the findings of view shift linear trends in regions outside of the VOT, it is necessary to consider the functional characteristics of those regions. The linear trend in voxels immediately posterior to the PPA is likely to reflect view sensitivity which would be predicted in posterior, place selective regions (Epstein et al., 2003; Epstein, 2005). Activation in V5/MT is thought to reflect motion selectivity driven by a greater percept of motion in the higher view shift conditions relative to the lower view shift conditions (Goebel, Khorram-Sefat, Muckli, Hacker, & Singer, 1998; Tootell et al., 1995). Successive images within the blocks were separated by fixation periods of 200ms, which was designed to reduce the percept of motion to achieve equivalence between same and different identity conditions. It has been suggested that cross talk between dorsal and ventral visual regions can allow the interpolation of unseen rotated views of objects (Weigelt, Kourtzi, Kohler, Singer, & Muckli, 2007). Such motion effects may be unavoidable when successive stimuli are processed as the same by the visual system, and motion processing in dorsal regions may be important for invariant recognition processes.

The linear trend across view shifts seen in the hippocampus has two possible interpretations. The first is that the activation reflects immediate visual short term memory, whereby the hippocampus stores a representation of the image such that comparisons with subsequent images within the block can be carried out. In the higher view shift conditions, successive images differ more than in lower view shift conditions, and as such could elicit higher activations in the hippocampus. Hartley et al (2007) found that patients with focal hippocampal damage showed impaired short term topographical memory (2s delayed match to sample) yet were unimpaired on non-spatial memory tasks. The present experimental stimuli were derived from the same test used by Hartley et al. which were explicitly designed to require hippocampal processing. These results are suggestive of hippocampal involvement in the immediate perception of the visual scene in order to build a coherent percept from a disjointed sequence of views.

Another interpretation of the pattern of results in the hippocampus is that the linear trend reflects adaptation of place cells involved in representing the virtual environment of the landscape stimuli. There is evidence of cells in the rat hippocampus that increase their firing rate when the animal moves through particular regions in the environment (Muller & Kubie, 1987; O'Keefe & Conway, 1978). Similar cells have also been found in human epileptic patients undergoing invasive monitoring with intracranial electrodes whilst they navigated virtual reality environments (Ekstrom et al., 2003). Thus far it has proven difficult to detect place cell activation in the human using fMRI. Although some limited success has been achieved using MVPA methods (Hassabis et al., 2009). In each of the view shift blocks within the present experiment, participants translate around the environment by parametrically varied amounts. In the lower view shift conditions, participants are moving very little or not at all, in the higher view shift conditions, participants are moving by larger distances. It has been suggested that the hippocampus represents spatial position with overlapping representations of particular locations within an overall map of the environment (O'Keefe & Nadel, 1978). It is plausible to suggest that this linear trend in activation within the

hippocampus is driven by the adaptation of adjacent, overlapping location representations within the hippocampal environment map.

### *2.4.6 Summary and Conclusions*

The present study was able to show that the response in the place selective regions is modulated by the degree of view shift encountered. Previous research was limited to testing hypotheses regarding the presence, or absence of view invariant representations. The parametric adaptation approach is able to go beyond this and describe the degree to which view invariance can be found in place selective brain regions and conclude that view invariance for scenes does indeed vary by degree rather than being all or none. The PPA adapts to view shifts of up to 10°, suggesting that the representation in this region is invariant to small changes in viewpoint but sensitive beyond this level. The partial view invariance in the PPA could provide a cohesive sense of location when naturally moving through the environment. Sensitivity in the PPA to larger changes suggests that additional processing must be needed for the degree of view shift encountered in the context of topographical memory and navigation. The present results indicate that such processes appear to involve the RSC which showed increased activation in individuals with better topographical memory. The pattern of adaptation shown in object selective regions suggests that object identity information may form part of a representation of scenes that subserve spatial navigation. The hippocampal response is modulated by the degree of view shift, perhaps driven by visual short term memory for the immediate scene, or alternatively by adaptation in hippocampal place cells sensitive to translational movement around the virtual landscape.

This chapter established that place selective regions showed degrees of viewpoint invariance, which in turn shed light on their respective roles in spatial behaviour. However, there was also some evidence of view-sensitivity for scene stimuli in area LOC which is not selective for such stimuli and typically considered to be specialised for object processing (Grill-Spector et al., 1999; Grill-Spector, Kushnir,

Hendler, & Malach, 2000; Malach et al., 1995). This raises the question of whether the mechanisms of view-invariance are specific to particular stimulus types or perhaps more general. If the former, a different pattern of adaptation throughout category selective cortex might be expected when participants process isolated objects seen from different viewpoints. This question forms a key motivation for the experiment described in chapter 3.

The present results indicate that the representation of landscapes in the PPA is sensitive to the degree of view shift encountered. It is possible that there are subpopulations within this region that show varying degrees of invariance. A key limitation of conventional ROI approaches is that they aggregate the response across all voxels in the region, limiting the ability to detect spatial variations in the responses. The present study also reported a whole brain linear trend measure, which is able to reveal voxels showing sensitivity to view shift. There is a difficulty with this approach in revealing voxels that show a completely view invariant patterns of response. It is possible to infer from the detection of place selective adaptation within the PPA, and the absence of voxels showing a linear trend across view shifts within the borders of this region, that the PPA is likely to contain invariant subpopulations. In order to show the extent to which this is true, and how invariance varies throughout place selective cortex, it is necessary to go beyond conventional analyses methods and combine independent measures of invariance and selectivity at the voxel level. This question is returned to, and addressed in detail, in chapter 5.

## **3 Processing of objects in object selective cortex: effects of parametric manipulation of viewpoint**

### **3.1 Introduction**

The present chapter focuses on the representation of objects within object selective cortex, and investigates the extent to which category selective VOT regions are invariant to changes in view. A key aim of this thesis is to investigate the nature of neural representation by measuring invariance to different kinds of stimulus change. Chapter 2 investigated this question for landscapes in place selective cortex, and found that view invariance in place selective cortex, when measured using a parametric manipulation of view shift, is dependent upon the degree of view change in the stimuli. The previous chapter also found that the response in object selective LOC was modulated by the view shift of landscapes. This raises the question of whether all category selective VOT regions show similar patterns of response to their non-preferred stimuli. Therefore the current experiment aims to investigate the neural representation of objects within the VOT and employs a similar experimental design to that of chapter 2, using a parametric manipulation of view shift applied to real world objects. This approach allows for comparison of the view sensitivity of category selective regions.

#### *3.1.1 Building invariant representations of objects*

In order to perform effective object recognition, it is necessary to build representations that are robust to the many kinds of change in the visual input that results from interacting with the world. The same object may be encountered close up, or far away; under different lighting conditions; it may be examined closely and directly, or briefly glanced out of the corner of the eye; or it may be encountered

from unique and previously unseen views. Each of these different kinds of change in the visual input will have an effect on recognition, however, under most circumstances the brain is able to recognise objects irrespective of any number of these transformations. It has been suggested that one of the key functions of the visual system is to facilitate this kind of recognition by transforming the retinal input into representations that are increasingly invariant to the different kinds of visual change (Dicarlo & Cox, 2007; Hung, Kreiman, Poggio, & DiCarlo, 2005).

There have been many computational theories of object recognition which have given specific descriptions of how different kinds of invariance are established. A key factor that differentiates many of these theories is the extent to which they postulate representations that are invariant to changes in view. Broadly speaking, there are two classes of theory, those that rely on three-dimensional object representations, and those that rely on viewpoint-dependent two-dimensional object representations. The former class posits that recognition takes place by fully reconstructing the 3d structure of a viewed object before comparing this to a stored 3d representation of that object (Marr & Nishihara, 1978). An extension of this theory which has received some empirical support, is the recognition by components theory (RBC), developed by Biederman (1987). RBC theory suggests that objects are represented in the form of 3D structural relationship between a restricted set of volumetric primitives known as “geons”. The latter, view-dependent class of theory suggests that object representations take the form of multiple 2D viewpoint-specific linked snapshots. And that recognition processes involve matching incoming images to stored views, or to transformed versions of stored views (Tarr & Pinker, 1989). This class of theory would suggest that fully 3D representations are not necessary for recognition and predicts that they would not be computed anywhere in the brain.

Of all the different forms of invariance that are needed for effective recognition, the most computationally difficult to establish is thought to be view invariance (Riesenhuber & Poggio, 2000; Ullman, 1989). Forms of invariance such as position or size, while the brain mechanisms that establish them are still not fully



understood, the mathematical operations necessary are simpler than those required for view invariance. In principle these types of invariance can be computed in a single step. View invariance is more likely than other forms of invariance to be built up in incremental steps, and vary by degree in any given brain region (Riesenhuber & Poggio, 2000). Furthermore it is possible that the problem of view invariance may be solved without ever establishing fully 3D representations (Ullman & Basri, 1991; Ullman, 1989). The choice of view shift as a manipulation is the most likely to yield a measure of invariance that will vary depending upon brain region, the selectivity of a given voxel or brain region, and the level of the visual processing hierarchy.

### *3.1.2 Visual object processing in the brain*

A great deal of research has been done looking at the nature of object representations in the brain, particularly the category selective VOT. The ventral processing stream feeds forward from low level visual areas to areas specialised for memory in the medial temporal lobes, and appears to be specialised for processing visual information in order to perform invariant recognition (Goodale & Milner, 1992). There is evidence that focal lesions of the occipital temporal regions result in associative agnosia specific to objects while sparing perception of other categories such as faces and scenes (Feinberg, Schindler, Ochoa, Kwan, & Farah, 1994). Neuroimaging studies have shown functional subdivisions within the VOT that show selectivity for objects as compared to scrambled objects (Malach et al., 1995). It is clear that this part of the brain is crucially involved in both object recognition and object perception, however, the nature of the representation in the object selective LOC remains an area of active research and debate.

There are several psychophysical studies which show that recognition speed is related to the rotational distance from the familiar view of the object (Bulthoff, Edelman, & Tarr, 1995; Tarr & Bulthoff, 1998). The dominant view of object recognition is that it is achieved entirely using view dependent processes. In

support of this, functional neuroimaging studies suggest that the LOC shows a view dependent pattern of response (Ewbank et al., 2005; Grill-Spector et al., 2001). There are however, conflicting accounts in the neuroimaging literature. Vuilleumier et al (2002), found that the LOC in the left hemisphere showed view invariant adaptation, but the right LOC did not. Furthermore there is evidence of functional subdivisions within the LOC whereby the more anterior posterior fusiform (pF) region, shows a greater degree of invariant adaptation than the posterior LO region (Grill-Spector et al., 1999; Kourtzi et al., 2003). Single cell recordings from inferotemporal cortex in monkeys, considered to be homologous to the VOT (Malach et al., 1995; Sawamura et al., 2005), have revealed neuronal populations which are invariant to changes in view and size (Booth & Rolls, 1998; Lueschow, Miller, & Desimone, 1994). Yet there have also been single-cell studies which have shown a greater proportion of view-dependent cells than invariant ones (Ashbridge, Perrett, Oram, & Jellema, 2000; Logothetis, Pauls, Bülthoff, & Poggio, 1994).

Whilst it is clear that the LOC is critically involved in object recognition and perception, there remains discrepancy about the extent to which this region is invariant or sensitive to changes in view. The different findings between previous studies may be the result of methodological differences. Experiments looking at single unit responses are limited by sampling of individual neurons, where it can be unclear how these responses summate at the population level. Many fMRI designs which investigate the issue of view invariance may be limited by the choice of stimuli and the size of the view shift manipulations used. The parametric adaptation approach can be used to characterise the underlying view invariance response in object selective regions.

### *3.1.3 View sensitivity throughout the VOT*

The previous chapter showed that both the PPA and the LOC are sensitive to the view shift of landscapes. The processing of scenes to facilitate recognition and spatial navigation requires that both spatial information about the scene, and the

spatial arrangement of the objects relative to the environment be encoded (Milner & Goodale, 1995). Object stimuli in isolation do not contain any information about spatial geometry of the environment therefore it is predicted that place selective regions will not be sensitive to view shift for objects. A parametric manipulation of view shift for objects comparable to that used for landscapes, considering the results of both studies, allows the relationship between category selectivity and invariance to be investigated.

A further reason to predict a dissociation in view sensitivity between the object and place areas and their respective preferred stimuli, is the large scale retinotopic eccentricity bias described in category selective VOT cortex (Levy et al., 2001; see chapter 1 for discussion). When viewed in an ecological setting, scene stimuli typically involve the entire visual field, whereas, for objects, it is common to position the eyes and head so that the object falls within central vision. Place selective activations are located in the more medial areas of the ventral surface of the brain, whereas object selective areas are found more laterally. This organisation predicts that place selective regions are unlikely to be sensitive to object view shift.

It was previously identified that image level differences between successive images in a block can vary systematically across conditions, as such it is possible that this low level confound could account for some of the modulatory effects of view shift (chapter 2). This was thought to be unlikely as a similar pattern of response was not seen in V1, which should be exquisitely sensitive to low level image changes. It is however, difficult to completely rule out the contribution of low level visual properties to the observed effects, and as such the present experiment includes a control condition whereby the objects are Fourier scrambled but the image level change between successive images within the block are equivalent to a 10° view shift. This control manipulation makes it possible to determine the extent to which adaptation at the level of the VOT may be driven purely by low level visual change.

### 3.1.4 *Summary and aims*

There are two main aims for the present chapter; to investigate the nature of the representation of objects within object selective cortex by measuring the extent to which these brain regions are invariant, or sensitive, to changes in view; and to investigate the extent to which view sensitivity differs throughout category selective cortex.

A crucial limitation of many previous studies of view invariance for objects is that they compare the response of a zero change condition to a view shift condition of some arbitrary amount, which may or may not be too large to elicit an adaptation effect, or show a view invariant pattern of response. View invariance may best be viewed as a continuum which is built up gradually throughout the ventral stream, with posterior regions invariant to small changes in view, and the degree of invariance increasing towards anterior regions (Grill-Spector et al., 1999; Kourtzi et al., 2003). If this were true, then the categorisation of a regional response as either view invariant or view dependent may depend upon the size of the view shift manipulation used.

A parametric manipulation of the view shift of object stimuli allows probing within the representation space of a given region in order to test the sensitivity of the underlying representation to the viewpoint dimension. Participants viewed blocks of images depicting real-world objects, the degree of view change between successive images within a block increased parametrically across conditions. When the degree of view shift within a condition is small, then the sampling distance along the view shift dimension is small. When the degree of view shift is large, distance along the view shift dimension is large. The degree of adaptation found within object selective cortex for a given level of view shift, reflects the sensitivity of the neural representation to the dimension of viewpoint.

In the present study a block design similar to that used by Ewbank et al.(2005) was carried out with the addition of a parametric manipulation of view shift (equivalent to chapter 2). Four view shift conditions were used ( $0^\circ$ ,  $5^\circ$ ,  $10^\circ$ ,

15°). If object selective regions such as the LOC are invariant to view shifts of the largest size (15°), and adaptation is an all or none phenomenon, then equivalent amounts of adaptation will be seen for all conditions. However, if object selective regions of cortex show an increase in response with increasing view shift, then the response will show a linear increase in response across view shift conditions. It is predicted that the response in the LOC will be modulated by the degree of view shift.

Ewbank et al previously found no view invariant adaptation for objects in any region using a view shift of 15°, although, both LOC and PPA showed view invariant adaptation to scenes at this level of view shift. In line with this, the findings of chapter 2 show that the pattern of response is modulated by the view shift of scenes for both of these regions. Object stimuli contain no environmental geometry, and there is evidence that this is necessary to drive place selective activity (Epstein, Harris, Stanley, & Kanwisher, 1999; Chapter 2). Additionally there is a large scale centre-periphery organisation throughout the VOT, whereby medial regions (such as the PPA) are driven by peripheral or full field stimuli (Levy et al., 2001). These two factors predict that the PPA response will not be modulated by the view shift of object stimuli, which are size matched and presented centrally in the fovea and without any environmental geometry.

## **3.2 Methods**

### *3.2.1 Participants*

22 healthy right handed participants took part in the study (7 male and 15 female) with a mean age of  $25.6 \pm 5.3$ ; all had normal, or corrected to normal vision, they were drawn mostly from the University of York student population. Written consent was obtained from all participants and the study was approved by the York Neuroimaging Centre Ethics Committee.

### *3.2.2 Imaging parameters*

All experiments were carried out using a General Electric 3 Tesla HD Excite MRI scanner at the York Neuroimaging Centre (YNiC) at the University of York. An 8 channel, phased-array head coil (GE, Milwaukee) tuned to 127.4 MHz was used to acquire MRI data. Functional images were obtained using a gradient-echo EPI sequence (TR 3000ms, TE 25.6ms, flip angle  $90^\circ$ , matrix 128 x 128, field of view 28.8cm) with 38 contiguous axial slices at a resolution of  $2.25 \times 2.25 \times 3$ mm. Functional images were centred over the occipital cortex in the plane of the temporal lobes. Whole brain T1 weighted structural scans were obtained at 1mm isotropic resolution (TR 7800ms, TE 2900ms, flip angle  $20^\circ$ , field of view 29cm). A T1-weighted flair image was also taken in the same plane as the EPI sequence to aid registration.

### *3.2.3 Object stimuli*

The stimuli used in the scanning task and subsequent memory test were drawn from the Amsterdam Library of Object Images (<http://staff.science.uva.nl/~aloi/>). 296 unique inanimate objects were manually selected to avoid unusually large, small or radially symmetrical stimuli and the presence of particular object categories such as faces or body parts (stimuli known to drive regionally selective responses in visual cortex). No object appeared in more than a single trial. The images were presented in colour on a black background and the depicted objects

were always centred within the visual field. In order to control for variation in object size, a measure of the pixel area for each object was taken and used to ensure the average visual field stimulation was equal across conditions.

### *3.2.4 Behavioural task*

Immediately after completion of the scanning task, participants were asked to complete a recognition memory task of objects they had just seen. The task was conducted using E-Prime (<http://www.pstnet.com/products/e-prime/>) and participants carried out a 64-trial, 4-alternative match to sample task. Success on the task depended on the formation of robust representations of objects encountered during scanning. Participants were aware that there would be a memory task following the scanning session. Each trial consisted of a display with four object images, one target previously viewed during the scanner task, and three unfamiliar foils. The task was to choose the previously viewed object. All targets were presented from previously unseen viewpoints. Participants were given a maximum of 20 seconds in which to respond.

### *3.2.5 Localiser scan*

Subjects viewed 20 blocks of 10 images. Each block contained images from one of five different categories: faces, bodies, objects, places, or Fourier scrambled images of the former categories. Face images were taken from the Psychological Image Collection at Stirling (PICS; <http://www.pics.psych.stir.ac.uk>) and body images were taken from a body image collection at Bangor (Downing et al. 2001; <http://www.bangor.ac.uk/~pss811/page7/page7.html>). Images of other categories were taken from a variety of web-based sources. Each image was presented for 700ms followed by a 200ms fixation cross. Stimulus blocks were separated by a 9s fixation screen. Each condition was repeated four times, and arranged in a counterbalanced block design.

### 3.2.6 *Adaptation scan*

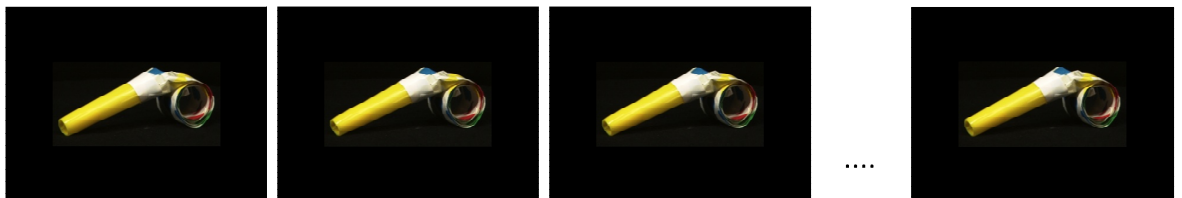
There were six experimental conditions with eight blocks for each condition. No object ever appeared in more than one condition. The six conditions were: (1) Different objects, random views selected from a pool of 104 objects; (2) Same image, same objects, 0° view shift; (3) Same object, 5° view-shift; (4) Same object, 10° view-shift; (5) Same object, 15° view-shift; (6) Scrambled object, image change within block equivalent to a 10° view-shift (**Figure 3.1**). Condition 6 is designed as a control for low level image change within an adaptation block. It is a scrambled version of the 10° shift condition, a random phase shift was generated for the first image within the block, then this phase shift was applied to the Fourier components of each successive image. As the images were in colour, this was applied to each colour channel of the RGB images. The degree of image change in the scrambled block was equivalent to the 10° condition but an intact object cannot be perceived. There were 48 blocks in total, which were sorted in a pseudo-random order, shuffled sequences of all six conditions occurring back to back throughout the functional run. Each image was presented for 800ms and separated by 200ms fixation cross. Each block consisted of 9 images, lasting 9s in total. The blocks were separated by fixation cross interval lasting 9s.



**Different objects:** random views



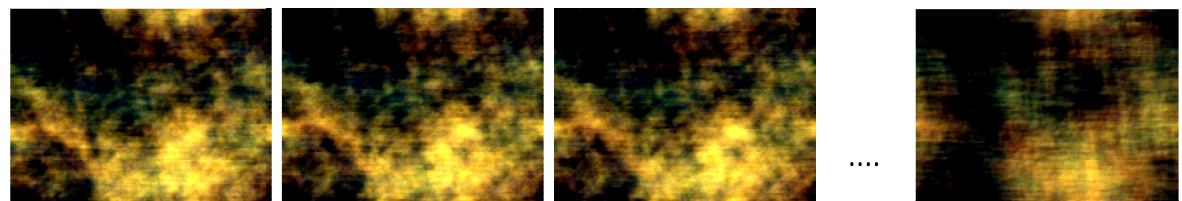
Same image: same object,  $0^\circ$  view shift



Same object,  $10^\circ$  view-shift, ordered sequence



Same object,  $10^\circ$  shift, phase preserved scrambled sequence



**Figure 3.1.** Examples of images used in the adaptation scan.

*Different objects (top row), each image within the block depicts a novel object from a randomly chosen view. Same object, same image (second row), a single view of a object repeated throughout the block. Same object,  $10^\circ$  view shift (third row), each image progresses around the object in an ordered sequence of  $10^\circ$  steps. Same object,  $10^\circ$  scrambled (bottom row), images from  $10^\circ$  shift conditions scrambled preserving Fourier phase shifts, the amount of change between successive images is equivalent to the  $10^\circ$  intact condition, designed as a control for low level image changes within a shift condition. Also included were  $5^\circ$  and  $15^\circ$  view shift blocks (not shown).*

### 3.2.7 fMRI attention task

In order to maintain attention throughout the experiment, participants were asked to respond to a red dot superimposed at a random location on the image. Either one or two dots were present in each block and the distribution of one dot and two

dot blocks was uniform across conditions. The red dot detection task was used in both the localiser and experimental scan. Results of the task were used primarily to verify attention and were not considered in the main analysis. An accuracy level lower than 80% on the red dot task was deemed to be an appropriate exclusion criteria, no participant's score fell below this threshold.

### 3.2.8 *fMRI analysis*

#### *Whole brain*

Whole brain analysis of the fMRI data was carried out using FEAT (<http://www.fmrib.ox.ac.uk/fsl>; Smith et al., 2004). The initial 9 s of data from each scan were removed to minimize the effects of magnetic saturation. Motion correction was done, along with removal of non-brain structures, spatial smoothing (Gaussian, FWHM 6mm) and temporal high-pass filtering (cut off, 0.01 Hz). A two-step mixed-effects analysis was carried out. An initial fixed effects model of each of the six conditions relative to inter-block fixation baseline was used to analyse individual participant data. Group analysis was carried out using a random-effects model (FLAME, <http://www.fmrib.ox.ac.uk/fsl>). Registration to high-resolution individual space and standard MNI space was done using FLIRT (FMRIB's Linear Image Registration Tool) for each individual (Jenkinson et al., 2002). Contrasts were calculated for each view shift condition relative to a baseline of the different object condition; significant activity in these contrasts is considered to reflect adaptation due to the same depicted object. A contrast of the 10° view shift condition with the scrambled object condition was calculated in order to show differences in response due to the presence of coherent view-shifting objects controlling for the influence of low level image characteristics. A contrast of all object blocks vs. fixation baseline was calculated in order to aggregate activity in all conditions and show regions selective for object stimuli. In order to take advantage of the parametric manipulation of view shift, a linear trend contrast was computed across all four view shift conditions (same image, same object condition was considered as a 0°

view-shift). The degree of significance within a voxel for this contrast reflects the extent to which activity increases linearly with increasing view shift. Such a contrast does not show activity relative to a baseline condition and as such is sensitive to trends across conditions occurring at all absolute activity levels. This contrast is able to detect differences which may not reach significance in conventional comparisons. Unless otherwise stated, whole brain images were thresholded at  $P < 0.001$  (uncorrected).

#### *Behavioural regressor*

Whole brain group analysis was repeated including a regressor derived from each individual's score on the post-scan recognition memory behavioural task. Individual scores were transformed into a deviation from overall mean score ( $(x^i - \text{mean}(x)) / \text{max}(x)$ ). Additional contrasts were added to reflect responses to each view shift condition relative to a baseline of grey screen fixation.

#### *Regions of interest*

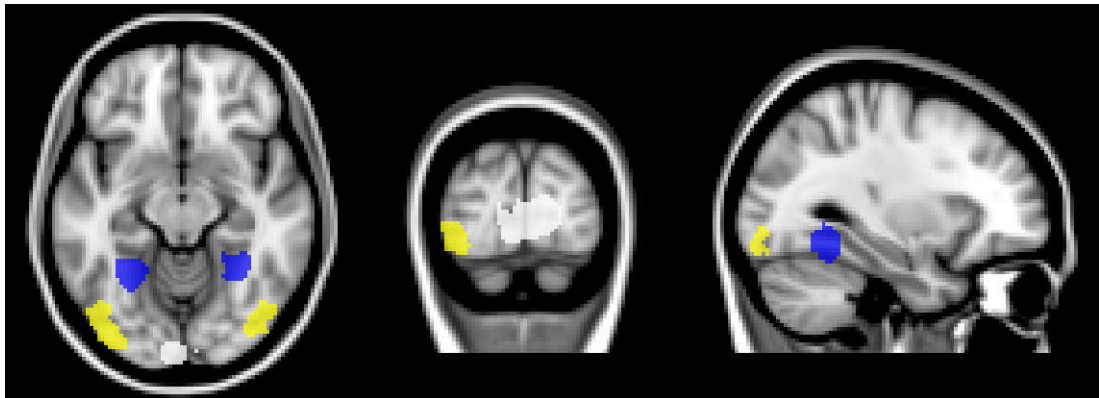
The contrasts of objects > scrambled objects from the localiser scan were used to define the LOC region of interest (ROI) in the adaptation scan, and places > faces was used to define the place selective PPA. All functionally localised regions were defined in each individual by anatomically constraining the activation using probability maps from the Harvard-Oxford cortical structural atlas (<http://www.fmrib.ox.ac.uk/fsl/fslview/atlas-descriptions.html>), and selecting up to the 50 most active voxels within this mask. A minimum threshold of  $P < 0.001$  (uncorrected) was applied, and individuals not showing activation above this threshold for a given region were excluded from the analysis. The V1 ROI was defined anatomically by masking the calcarine sulcus of each individual structural scan, and further constraining the mask using active > rest in the localiser scan, ensuring that selected voxels were responsive to the area of the visual field stimulated throughout the experiment.

GLM parameter estimates for each adaptation scan condition were extracted for each ROI and all voxels within each region were averaged together. Activation values for each region were converted from arbitrary units to units of % signal change relative to fixation baseline ( $\text{parameter estimate}/100 \times \text{Regressor Height}$ ). Responses from all ROI's were subjected to repeated measures ANOVA to determine significant differences between stimulus conditions, regions and hemispheres.

### 3.3 Results

#### 3.3.1 Localiser

Object and place selective regions of interest were identified within the (VOT) region of the brain (**Figure 3.2; Table 3.1**). The LOC was identified as a region in the lateral occipital cortex or posterior cingulate showing a greater selectivity for objects relative to scrambled objects (Malach et al., 1995). The PPA was identified as a region along the parahippocampal gyrus that showed a greater selectivity for place stimuli relative to face stimuli (Epstein & Kanwisher, 1998). Both place and object contrasts elicited sizable bilateral regions of activation in all participants. The V1 ROI was created anatomically by manually identifying the calcarine sulcus in each individual and constraining the region using a functional contrast of active>rest in the localiser scan. This ensured that only active portions of the visual field were sampled. Masks were created for each of the above regions which were then used to extract mean time series' for each region within the adaptation scan for all six conditions.



**Figure 3.2.** Category selective, localiser scan ROI masks shown in MNI space. Yellow=LOC (objects>scrambled), blue=PPA (places>faces), white=V1 (anatomical calcarine sulcus combine with active>rest). Masks are created using a minimum threshold of  $P < 0.001$  (uncorrected). Images are centred on MNI coordinates: -30, -90, -12. Scans are displayed using neurological convention with the right hemisphere displayed on the right.

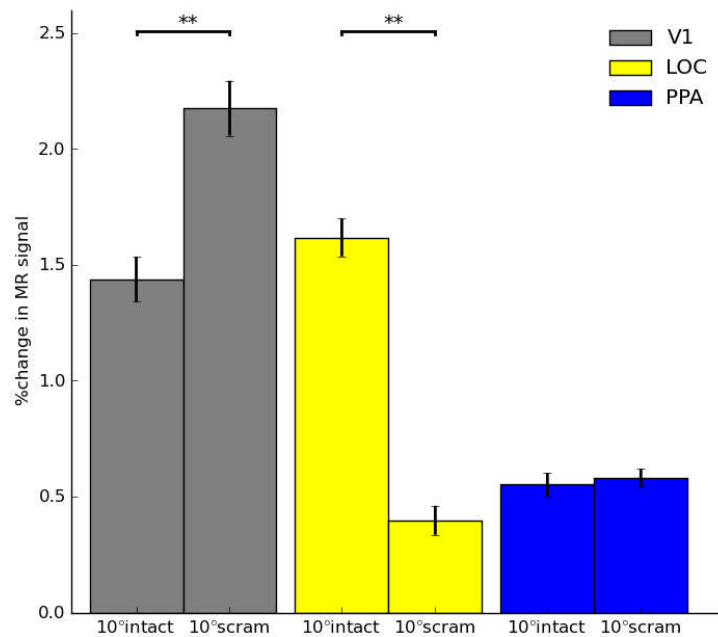
**Table 3.1.** MNI coordinates of regions of interest peaks

Area	Hemisphere	MNI Co-ordinates			Z-value
		X	Y	Z	
LOC	Right	34	-86	-2	6.48
	Left	-40	-68	-18	5.92
PPA	Right	24	-50	-16	5.91
	Left	-26	-48	-16	6.28
V1	Right	14	-94	2	7.09
	Left	-10	-98	-12	6.46

### 3.3.2 Scrambled control maintaining visual change

It is possible that any effects that are observed in response to intact view shift conditions result from variation in terms of low level image features, which increase, and are potentially confounded with increasing view shift. For this reason a control condition was included consisting of a Fourier scrambled version of the 10° view shift condition, well controlled in terms of low level featural differences between successive images within a block. Image order was maintained such that the degree of low level visual change between each successive image remained consistent with the intact 10° condition.

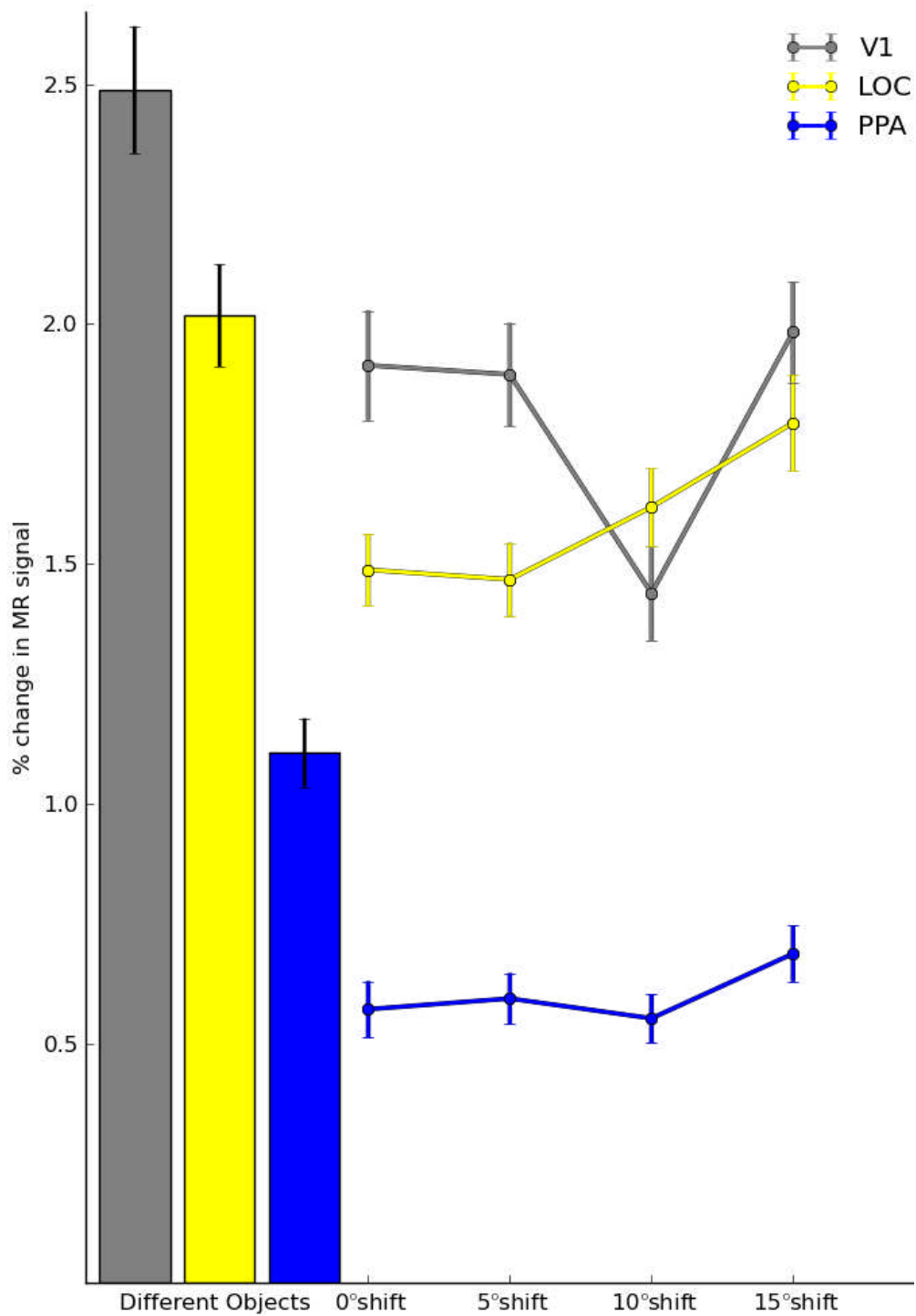
A whole brain comparison of intact with scrambled 10° conditions was completely consistent with the borders of the object selective LOC (objects>scrambled). Activation for the 10° scrambled condition was largely restricted to low level visual regions with higher level activation extending only into the medial portion of the ventral surface. A t-test was conducted to compare the intact and scrambled 10° conditions in each of the ROIs. The LOC shows significantly greater activation to intact ( $M = 1.619$ ,  $SE = 0.83$ ) compared to scrambled ( $M = 0.398$ ,  $SE = 0.064$ ) shift conditions,  $t^{(21)} = 18.916$ ,  $p < 0.001$ . V1 showed a greater response to scrambled ( $M = 2.175$ ,  $SE = 0.099$ ) compared to intact ( $M = 1.439$ ,  $SE = 0.121$ ) shift conditions,  $t^{(21)} = -8.108$ ,  $p < 0.001$ . The PPA showed almost identical responses to intact and scram shift conditions. See **Figure 3.3**.



**Figure 3.3.** ROI responses to intact and scrambled 10° shift conditions. Error bars represent  $\pm 1$  standard error of the mean. \*\* indicate statistical significance at  $p < 0.001$  (uncorrected).

### 3.3.3 Adaptation

No main effects or interactions involving hemisphere were found in any ROI, responses across hemispheres were pooled for all subsequent analysis. The responses for all of the view shift conditions relative to the different condition can be seen in **Figure 3.4**. Planned adaptation comparisons for different objects compared to the same objects conditions were carried out for all regions (**Table 3.2**).



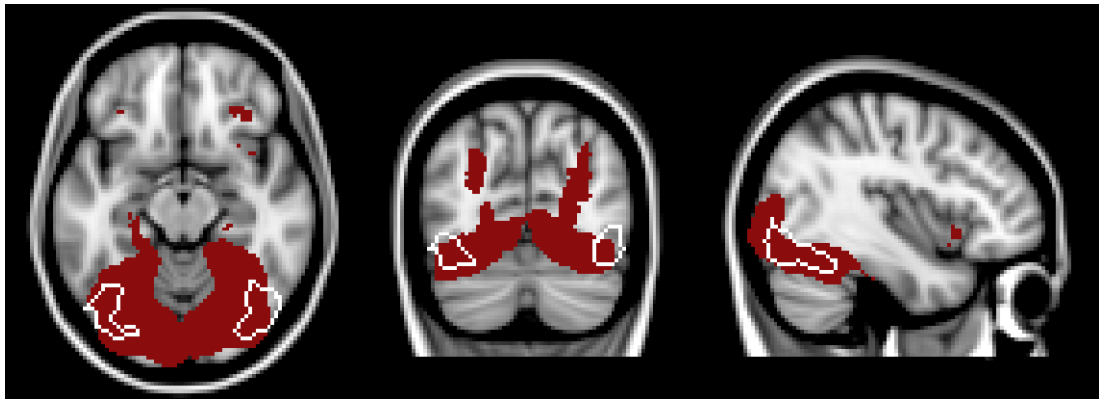
**Figure 3.4.** Response to all view shift conditions in all ROIs.

Bars represent the response to different objects condition; connected line represents the response across view shift conditions. Data are shown for LOC (yellow), PPA (blue), and V1 (grey). Significant linear trend across view shift conditions observed only in LOC. Error bars represent  $\pm 1$  standard error of the mean.



**Table 3.2.** *Different-Same adaptation contrasts, all ROIs.*

ROI	Adaptation Contrast:- Different - Same (0°,5°,10°,15°)
LOC	$F_{1,21} = 68.628, P < 0.001$
PPA	$F_{1,21} = 157.139, P < 0.001$
V1	$F_{1,21} = 158.789, P < 0.001$

**Figure 3.5.** *Whole brain same object adaptation.*

Areas in red are voxels showing adaptation to the same object regardless of view shift, as defined by the contrast different objects > same object (0°, 5°, 10°, & 15° view shift conditions). Outlines of the LOC as defined by a contrast of objects > scrambled are shown in white. Statistical images are created using a threshold of  $P < 0.001$  (uncorrected). Images are centred on MNI coordinates: 40, -72, -14. Scans are displayed using neurological convention with the right hemisphere displayed on the right.

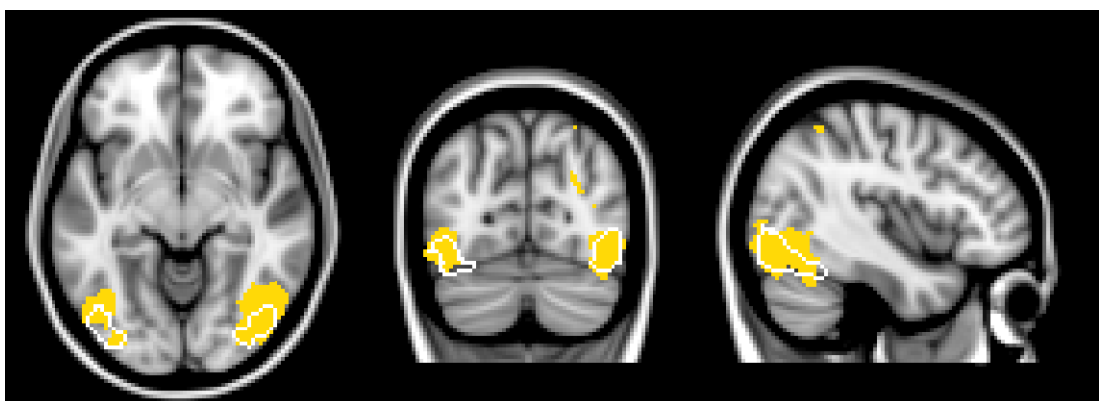
Substantial same identity adaptation was found in all ROI's. In the whole brain, almost all visually responsive regions showed significant same identity adaptation. **Figure 3.5** shows the whole brain map of regions responding more to different objects than same objects regardless of view shift. All contrasts of individual view shift conditions with different objects showed extensive adaptation, and as such, it was not possible to examine differences in adaptation response across view shift conditions.

### 3.3.4 View shift sensitivity

Objects depicted in all view shift conditions maintained the same identity throughout the block, yet these conditions differ only in the degree of view shift

between successive images. It was thought that a linear increase in view shift would result in a similar linear increase in BOLD response in object selective regions, but not in non-object selective regions. In order to look for differences in view shift response across all ROIs, a 3x4 repeated measures ANOVA was carried out (region, view shift). There was a strong main effect of region ( $F^{(2,42)} = 98.657, P < 0.001$ ), a main effect of view shift ( $F^{(3,63)} = 19.426, P < 0.001$ ). There was an interaction of region with view shift ( $F^{(6,126)} = 50.491, P < 0.001$ ). The same image condition was treated as a 0° view shift and a linear trend contrast across the four view shift conditions was computed. An overall linear trend main effect was observed across all ROI's ( $F^{(1,21)} = 9.349, P < 0.01$ ). Linear trend did however show an interaction with region ( $F^{(1,21)} = 61.521, P < 0.001$ ). When linear trend is examined independently for each region, there is no significant trend in V1 ( $F_{1,21} = 2.401, P = 0.136$ ), a weak but significant linear trend is observed in the PPA ( $F_{1,21} = 5.922, P < 0.05$ ), and a large significant linear trend is observed in LOC ( $F_{1,21} = 37.543, P < 0.001$ ). See the line graph section of **figure 3** for an illustration of these effects.

A whole brain linear trend contrast was carried out in order to identify regions that show a linear increase in activation with parametric increases in view shift. Activation in the linear trend contrast was found to correspond closely to the boundaries of object selective cortex and adjacent areas (**Figure 3.6**).

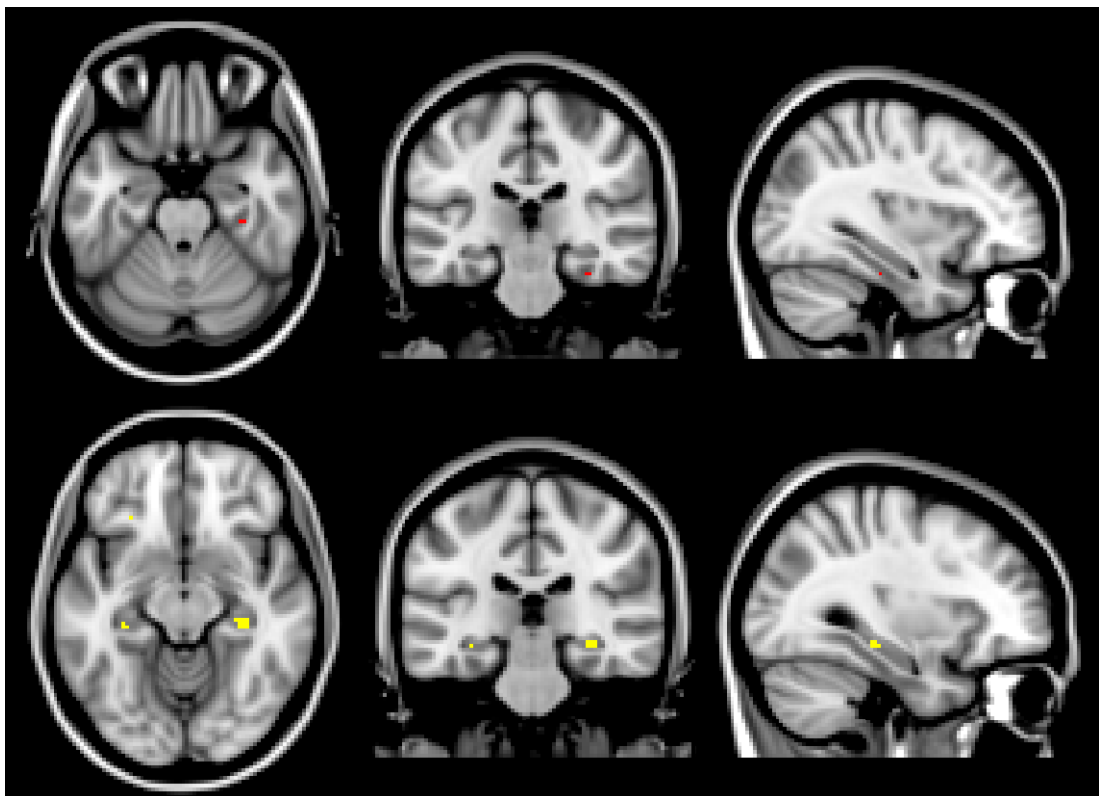


**Figure 3.6.** *Whole brain linear responses to view shift. Bilateral activation in object selective cortex that shows a linear increase in activation with successively increasing view shifts (yellow). This contrast was carried out using the 0°, 5°, 10°, & 15° conditions with a contrast vector -3, -1, 1, 3. Outlines*

of the LOC as defined by a contrast of objects > scrambled are shown in white. Statistical images are created using a threshold of  $P < 0.001$  (uncorrected). Images are centred on MNI coordinates: 42, -74, -8. Scans are displayed using neurological convention with the right hemisphere displayed on the right.

### 3.3.5 Behavioural measures

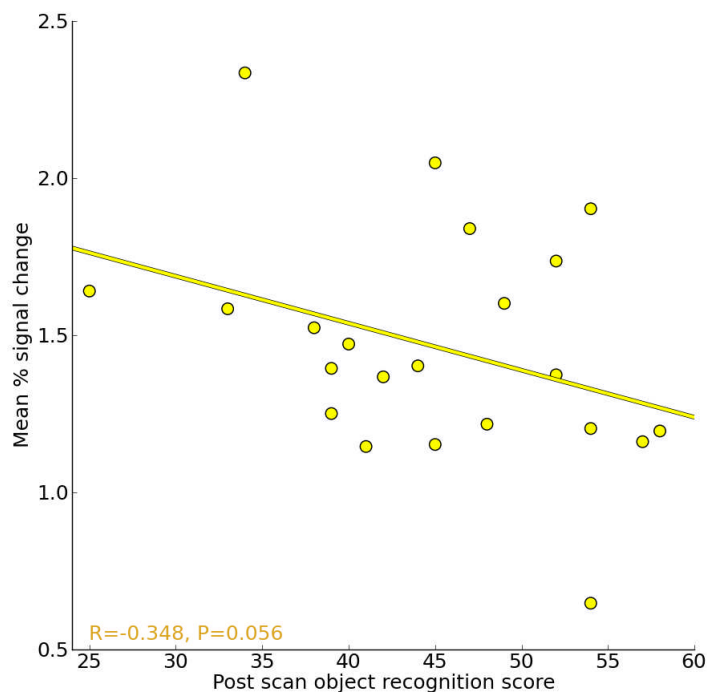
The average score in the post scan memory test (out of a maximum of 64) was  $45 \pm 8.5$  with most participants performing above chance. The minimum score was 25 and the maximum was 58. Scores were transformed into a deviation from the mean score and entered as a behavioural regressor into the whole brain analysis (see **methods 3.2.8**). Contrasts of the all stimulus blocks vs. fixation baseline were examined for behavioural correlations. Positive correlations between neural activity when viewing object stimuli and object recognition memory performance were observed in the right perirhinal cortex, and negative correlations for the two variables were found bilaterally in the medial hippocampus (**Figure 3.7**).



**Figure 3.7.** Whole brain correlations with object recognition memory. (Upper slices) Areas in red are voxels in the right perirhinal cortex showing positive correlation of object recognition memory performance and neural activity (active > rest in adaptation scan). (Lower slices) Areas in yellow are voxels in the

*middle portion of the hippocampus showing negative correlation of object recognition memory performance and neural activity (active>rest in adaptation scan). Statistical images are created using a threshold of  $P < 0.001$  (uncorrected). Images are displayed on the MNI152 standard brain, centred on MNI coordinates: 34, -26, -24 and 32, -28, -12 respectively. Scans are displayed using neurological convention with the right hemisphere displayed on the right.*

All ROIs were examined for correlations of neural activity with object recognition memory performance. Only the LOC showed negative correlation approaching significance (see **Figure 3.8**). This relationship was not observable in the whole brain analysis of the relationship between brain response and behavioural performance.



**Figure 3.8.** LOC ROI correlation with object recognition memory. Scatter plot showing the relationship between object recognition memory performance and BOLD response for all object conditions relative to fixation baseline.

### **3.4 Discussion**

When view shift is varied for object stimuli, the response in object selective cortex is modulated by the degree of view shift applied to the stimuli. This is evidenced by strong linear trends across the four levels of view shift throughout object selective cortex, as well as in the LOC ROI. No other region showed the same modulation of response when object view shift was varied. The primary aim of this chapter was to test the extent to which the representation in LOC is sensitive to changes in view. These results are consistent with previous fMR-a results which show that the LOC is view dependent (Ewbank et al., 2005; Grill-Spector et al., 1999, 2001), yet emphasise the importance of considering the size of view shift manipulations in interpreting such fMR-a results.

A further aim was to consider the extent to which non-object selective regions are sensitive to the view shift of object stimuli. As predicted, the responses in non-object selective regions were not modulated by view shift of object stimuli. The PPA did show a weak linear trend across the four view shift conditions but this pattern was completely absent in the whole brain analysis and the pattern of response was much flatter compared to that of the LOC, with this difference shown to be statistically reliable. The pattern of results of this chapter and those of chapter 2 are generally in agreement with the findings of Ewbank et al. (2005), who showed view invariant adaptation for scenes in the LOC and PPA but failed to show similar adaptation in those regions for similar manipulations of objects. Ewbank et al., used a view shift of 15°, which is equivalent to the highest view shift used here, the linear increase in response for increasing view shifts indicates that a view shift of 15° is close to the point where the response in the LOC can no longer show adaptation. At view shifts larger than 15°, successive images of objects are represented in the LOC as images of distinct objects.

*3.4.1 The effects of low level visual change*

It was thought that low level image changes could contribute to the pattern of responses observed using the parametric adaptation paradigm. In order to investigate this, a scrambled control condition was included in which the low level image properties varied to the same extent as the 10° intact condition. The condition was created by Fourier scrambling the images in the each of the 10° blocks, maintaining the phase between each successive image. The pattern of response for the 10° scrambled control condition compared to the 10° intact condition indicate that the adaptation effects seen across the view shift conditions are unlikely to be accounted for by low level image differences alone. The greater response for the scrambled condition in early visual regions is explainable as a result of the retinotopic extent of the visual stimulation in the 10° scrambled condition. The objects before scrambling occupied a small part of the central visual field, yet the Fourier scrambled images extended further into the periphery. The V1 ROI, which was likely to include voxels within the peripheral representation of the calcarine sulcus, showed a greater response when aggregated across all voxels in the region. The difference in visual field extent is also likely to explain the tendency of the scrambled condition to activate medial portions of the VOT, as these regions have been associated with a peripheral representation of the visual field in a coarse eccentricity bias known to be present in the VOT (Levy et al., 2001).

The pattern of response in primary visual cortex was expected to be largely flat, and not to vary between the four view shift conditions, yet the response in the V1 ROI showed a substantial reduction for the 10° shift condition which was not apparent for any other condition. This pattern is difficult to explain in terms of either low or high level features of the stimuli, as by any measure of differences between the shift conditions, the 10° condition falls between the 5° and the 15° conditions. The pattern of reduced response to this condition is observable at the whole brain level, and in the ROI responses for each individual (see **Appendix B. 1**), and as such appears to be a real effect that cannot be explained by outliers or

individual differences. It is possible that this condition differed in some systematic way from the others as a result of the objects that were randomly assigned to it. However this is thought to be unlikely as the effect is present for this condition only in V1 and not in any other region, and the images were chosen to be largely equivalent in terms of size and image features. An alternative explanation for this effect could be long-lag adaptation for low level properties caused by exposure to the 10° scrambled blocks, which were constructed by scrambling images in the 10° intact condition. All images were presented in colour and the 10° intact and scrambled conditions were well matched in terms of their global luminance and hue properties. It is possible that these low level similarities resulted in long-term adaptation in V1 for the scrambled condition and a resulting reduction of average response to the intact condition. V1 has been shown to be sensitive to low level properties such as hue and saturation (Hanazawa, Komatsu, & Murakami, 2000; Johnson, Hawken, & Shapley, 2008), as well as relative luminance (Peng & Van Essen, 2005). There is evidence that the timescale upon which adaptation operates varies depending on cortical area, with primary visual cortex sensitive to more long term adaptation manipulations relative to extrastriate regions (Fang et al., 2005).

### *3.4.2 Links between object recognition memory and brain response*

Participants were given a post scan recognition test for the objects that were viewed in the scanner. There were two regions in the brain that showed a relationship between recognition performance and brain response, both showing a negative relationship, these were the LOC, and bilateral regions in the medial hippocampus. The more active these regions were, the worse an individual performs at object recognition memory. Participants were explicitly aware that a memory test followed scanning and as such, may have engaged in various strategies in order to remember the items in the scanner. It has been suggested that the role of the hippocampus and surrounding structures in memory processes can be divided into separate sub-systems (Aggleton & Brown, 1999). This dual process

theory suggests that hippocampal-anterior thalamic regions subserve encoding and subsequent recall of episodic memory, whereas familiarity judgements do not require these systems and instead rely on the perirhinal cortex. The present findings fit well with this theory; participants who attempted to form declarative memories of each of the items they viewed in the scanner, engaging hippocampal systems, did worse in the recognition task. Participants who chose a passive viewing strategy, engaging perirhinal cortex, and familiarity processes, were better performers. The relationship observed in the LOC was a marginal effect, and was not observable in the whole brain even at relaxed threshold; therefore this effect is interpreted with caution. It could be speculated that attempts to use a declarative memory strategy results in increased attention to shape properties of the object stimuli, resulting in increased activation of the LOC, as well as limiting subsequent recognition memory performance.

### *3.4.3 Effects of attention*

The relative difference in BOLD response between different and each of the same identity conditions observed throughout all visually responsive brain regions was considerably higher than would have been expected from similar block design studies with equivalent comparisons (Ewbank et al., 2005). It is thought that the participants' understanding of a forthcoming memory test led to them attempting to remember all items viewed in the scanner, which was considerably more taxing during blocks of different objects containing nine rapidly presented objects. The different objects condition is thought to have involved increased attentional load and greater task difficulty, which could account for the unusually high BOLD response seen here. This leads to difficulty in interpreting the results of adaptation contrasts using the different condition as a baseline, as all comparisons produce highly significant differences. All of the same identity conditions, which do show differences in response across levels of view shift and brain regions, are well matched in terms of attention and task difficulty. It is likely that any differences



observed across the same identity conditions are unaffected by this issue. Any modulation of the response across view shift conditions can be interpreted as an effect driven by neural adaptation.

This finding does highlight the difficulty posed by a common adaptation baseline condition used across multiple conditions. Any feature of such a common condition which leads to an over- or under-estimation of its response can heavily influence the interpretation of overall patterns of response. In future studies of parametric adaptation, each level of a manipulation would ideally have a discrete independent baseline. Such a baseline would be problematic for manipulations of viewpoint, as sequences of different objects or landscapes do not necessarily have a common reference frame around which to shift in view. Such manipulations are possible however, for other kinds of invariance that the visual system is known to establish, such as size or position. Furthermore, this finding suggests that it would be worthwhile in future studies to tightly control subjects' attention across conditions, ensuring that the magnitude of the response measured for each condition is purely stimulus driven. Both of these factors are incorporated into the design of chapter 4, when considering a parametric manipulation of size change for objects.

### *3.4.4 Summary and Conclusions*

The present experiment was able to show that the degree of view invariance for objects found in object selective cortex is dependent on the size of view shift applied to the stimuli being viewed. A similar modulation of invariance by degree of view shift was observed for landscape stimuli, and found in both place selective and non-place selective cortical regions (chapter 2). For object stimuli however, this effect was confined to object selective cortex, did not apply to other category selective regions, and could not be accounted for by low level visual changes (chapter 3).

There is a dissociation in how viewpoint is processed for object and landscapes stimuli. Both object and place selective regions within the VOT are sensitive to the view shift of landscapes, whereas only object selective regions are sensitive to view shift for objects. This could be a reflection of the coarse retinotopic organisation of the VOT (Levy et al., 2001), or it could reflect separate streams of visual processing within the VOT for establishing invariant representations of objects and landscapes. Chapter 5 combines measures of view invariance derived from experimental chapters 2 and 3, as well as selectivity measures derived from localiser data, and investigates how the processes of invariance change spatially throughout the category selective cortex.

## **4 Parametric adaptation for size change in object selective cortex**

### **4.1 Introduction**

Experiments in previous chapters were primarily designed to quantify view-invariance in order to examine the nature of representations within and between category selective regions of cortex. It has been shown that the parametric adaptation paradigm is an effective way to measure view-invariance in category selective cortex, showing how large scale retinotopic organising principles of the VOT interact with view shift sensitivity in category selective functional regions. Both the PPA and the LOC show view sensitivity to their preferred categories, and the LOC additionally shows view sensitivity to places, the PPA however, showed no view sensitivity for objects. The PPA and the LOC regions therefore, can be dissociated on the basis of view sensitivity to their non-preferred stimulus.

The previous two experiments did not show completely invariant patterns of response in any extrastriate visual region. This may imply that truly view invariant representations are not needed to perform object or scene recognition, and as such would not be detectable anywhere in the brain (Tarr & Pinker, 1989). However, it is also possible that view invariant representations are established later, in more anterior, medial temporal lobe structures. Computational theories of object recognition have principally been concerned with the extent to which such processes require invariance to changes in view, and this is largely because there are many competing accounts of how the visual hierarchy might go about solving this particular kind of invariance (see chapter 3 for discussion). It is however, usually agreed that invariance to size change is a computationally easier problem to solve, that it must be solved in order to perform invariant object recognition, and that it is likely established early on in the processing stream. It is generally thought that

representations are built in stages moving up through the visual processing hierarchy (Ullman & Soloviev, 1999). As receptive field size of successive neural populations increase, neural responses become increasingly invariant to changes in position and size. Such increasingly invariant cells are receiving input from increasingly large sections of the visual field which may encompass the entirety of an object to be recognised.

The present study focuses on invariance to the size of objects in the visual field because at the level of the VOT cortex, receptive field size is often sufficiently large to encompass large portions of the visual field (Gross et al., 1972); and there is evidence from the fMRI literature, that by this stage, complete size invariance for objects has been established (Ewbank et al., 2005; Grill-Spector & Malach, 2001; Grill-Spector et al., 1999; Sawamura et al., 2005). Both experimental methodologies and manipulations of object size vary between studies, and as such it is unclear if complete size invariance is established at the level of the VOT/LOC, and under what circumstances.

#### *4.1.1 Size invariance in object selective cortex*

There have been several monkey single unit studies investigating size invariant processing in the IT region. Invariance in single unit studies is typically measured by presenting stimuli for which a cell shows selective stimulus driven activation, and testing for robustness of that response to image changes. Lueschow et al. (1994) recorded from cells during a delayed match to sample task with pictures of complex objects. The monkeys were trained to perform this task ignoring image transformations such as size, in which the image could appear at either  $2^{\circ 2}$  of visual angle, or  $4^{\circ 2}$ , or a position, in which the image appeared either foveally, or  $5^{\circ}$  into the contralateral visual hemifield from the recording site. During perception of the stimuli, the stimulus preferences measured by neural adaptation in individual cells was unaltered by size and position changes. However, some cells showed a preference for stimuli of a particular size or location. During the memory

component of the match to sample task, the majority of cells showed complete invariance to size changes. This study shows that there are cells within IT that are both size invariant and size dependent, suggesting that IT is directly involved in invariant recognition processes, whilst also mediating the perception of size and location changes of visual objects.

Ito et al. (1995) investigated invariance across a large range of sizes in shape selective regions and found evidence that both size invariant, and size dependent processing occur in the anterior most portion of IT (TE). They, and found that 21% of neurons studied were invariant to changes in size of up to 4x, whereas 43% of neurons showed some sensitivity in their responses to size change.

It has been shown that more anterior IT populations have larger receptive fields than those in the posterior part (Gross et al., 1972). The consensus from the single unit literature is that only a subset of IT neurons show invariance to size changes, and frequently the majority of cells recorded from show size dependent responses. This is a finding that has been found repeatedly across a range of similar studies (Miyashita & Chang, 1988; Schwartz, Desimone, Albright, & Gross, 1983). The more anterior region of the IT, area TE, shows a greater degree of size invariant responses than posterior parts of IT, which is in line with the view of a hierarchically organised processing stream with increasing selectivity and invariance at successive stages.

Single unit studies, whilst considered to be the gold standard method for measuring invariance, are always limited by the amount of neurons they are able to sample. Often sampling hundreds from a population of millions, and the location of sampling will differ slightly depending on the idiosyncrasies of a given animals brain anatomy. The method of fMR-adaptation is thought to be able to characterise the functional properties or neural populations at resolutions below that of conventional fMRI (Grill-Spector & Malach, 2001; see chapter 1 for discussion). FMR-a has been used extensively to investigate the issue of size invariance in object selective brain regions in the human. At the population level, the majority of studies

suggest that the LOC does show size invariant adaptation, which is somewhat at odds with the general findings of single unit studies.

Grill-Spector and Malach (2001) tested whether the fMR-adaptation responses in LOC were robust to changes in object size, position, illumination and view point. They found that LOC is invariant to changes in size and position in the visual field, but is sensitive to changes of illumination and viewpoint. Ewbank et al., (2005), conducted an investigation of view and size invariance in several VOT regions. They found that both the LOC and the PPA showed adaptation that was invariant to size changes ranging from  $3^{\circ 2}$  to  $9^{\circ 2}$ .

Grill-Spector et al., (1999) conducted an event related fMR-adaptation experiment in which participants were shown object stimuli which underwent various viewing condition changes, such as size, position, orientation and lighting conditions. Differences in the nature of the representation were found between the posterior and anterior portions of the LOC, leading the researchers to propose functional subdivisions of the region. The posterior region (LO) showed a release from adaption under all image transformation and in general, showed less invariance. Whereas the more anterior posterior fusiform (pF) showed greater invariance, displaying sustained adaptation to both changes in size and orientation.

Sawamura et al., (2005) found the same posterior/anterior differences in size invariance within object selective cortex as Grill-Spector et al. whilst also comparing responses in the human and the macaque. They found the same pattern in both species, confirming both the homology between the human LOC and the macaque IT, as well as the effectiveness of fMR-adaptation in studying invariance at the level of the neural population.

The fMR-adaptation literature suggests that complete size invariance for objects is established within the LOC (Ewbank et al., 2005; Grill-Spector & Malach, 2001; Grill-Spector et al., 1999; Sawamura et al., 2005). However, a limitation of the majority of previous experiments is that they make a single comparison of a size changing condition with a non-size changing condition, limiting findings to either the detection of size invariant adaption, or the absence of it. There are often large

differences in the degree of size change used between studies, and this could lead to differences in whether or not invariance is detected.

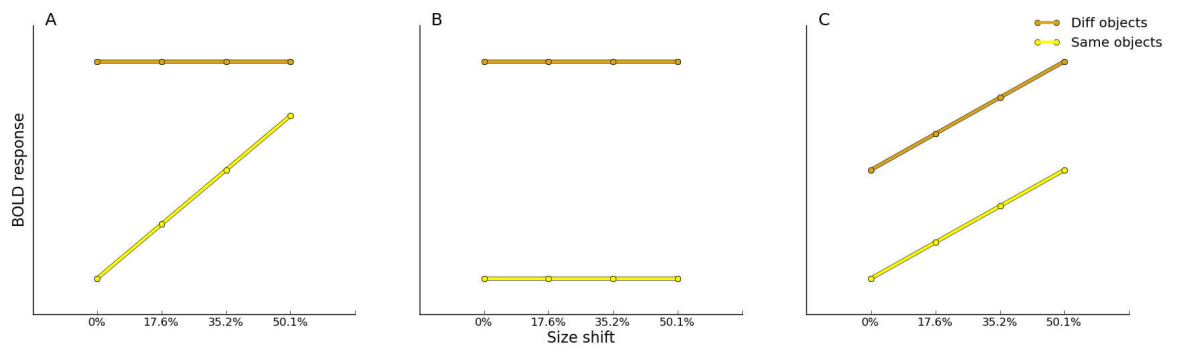
In this study, a parametric adaptation paradigm is applied to size change for real world, recognisable, and manipulable objects. There is a clear need for everyday objects to be recognised from both close up and far away, as well as a related requirement for information about the real-world size of objects to be calculated so that they can be manipulated and interacted with. This paradigm has previously been applied to view shift for objects and landscapes (chapters 2 & 3), and a limitation of the paradigm when applied to view shift is that it cannot be equally applied to different identity objects or scenes in the same way as it can be applied to same identity stimuli. When identity is varied, there is no common reference frame for each successive image and as such the viewpoint in different identity conditions is irrelevant. This limitation does not apply to manipulations of size however; change in the size of the visual field can be applied to both same- and different- identity objects without appreciably affecting recognisability. A parametric manipulation of size shift therefore allows an independent adaptation baseline condition for each level of shift.

### *4.1.2 Summary and aims*

A limitation of previous experiments investigating size invariance is that they make a single comparison of a size changing condition with a non-size changing condition, limiting findings to either the detection of size invariant adaption, or the absence of it. There are often large differences in the degree of size change used between studies, which may also lead to discrepant findings. The present experiment systematically investigates the extent to which object selective, and non-object selective, visual areas are invariant to changes in the size of object stimuli using a parametric manipulation of size change to both same- and different- identity objects. Furthermore, the range of visual field stimulation is held constant across all conditions. It is expected that if the object representation is sensitive to object size

change, then the degree of adaptation will be modulated by parametrically increasing size change, with higher size change conditions showing progressively less adaptation in larger size change conditions. This will not be the case if the representation is invariant to size change. It is possible therefore, to distinguish between size dependent and size invariant representations depending upon the profile of adaptation across the parametrically increasing size change conditions.

**Figure 4.1** shows the hypothetical responses for size dependant (a) and size invariant (b) representations, and also an additional possibility of a partially size sensitive, object selective response (c).



**Figure 4.1.** Hypothetical responses for parametrically size shifting object stimuli. (a), size sensitive, object selective response. (b), size invariant, object selective. (c), size sensitive, partially object selective.



## **4.2 Methods**

### *4.2.1 Participants*

26 healthy right handed participants took part in the study (9 male and 17 female) with a mean age of  $26\pm 6.9$ ; all had normal, or corrected to normal vision, they were drawn mostly from the University of York postgraduate population. Written consent was obtained from all participants and the study was approved by the York Neuroimaging Centre Ethics Committee.

### *4.2.2 Imaging parameters*

All experiments were carried out using a General Electric 3 Tesla HD Excite MRI scanner at the York Neuroimaging Centre (YNIC) at the University of York. An 8 channel, phased-array head coil (GE, Milwaukee) tuned to 127.4 MHz was used to acquire MRI data. Functional images were obtained using a gradient-echo EPI sequence (TR 3000ms, TE 25.6ms, flip angle  $90^\circ$ , matrix 128 x 128, field of view 28.8cm) with 38 contiguous axial slices at a resolution of 2.25x2.25x3mm. Functional images were centred over the occipital cortex in the plane of the longitudinal axis of the temporal lobes. Whole brain T1 weighted structural scans were obtained at 1mm isotropic resolution (TR 7800ms, TE 2900ms, flip angle  $20^\circ$ , field of view 29cm). A T1-weighted flair image was also taken in the same plane as the EPI sequence to aid registration.

### *4.2.3 Object stimuli*

Throughout the object adaptation experiment, a limited selection of objects were used in order to ensure that differences between conditions cannot be accounted for in terms of image level, or featural differences between the sets of objects in each condition. Eight objects were selected from the NOVA Development, Art explosion library of Object Images (<http://www.novadevelopment.com>). Objects were selected on the basis of recognisability, avoiding unusually asymmetrical objects, as well as avoiding images containing object categories such as faces or

body parts (i.e., stimuli known to drive regionally selective responses in visual cortex). Objects selected were all chosen to be of approximately the same real-world size and that were readily manipulable. The foreground area of each of the images was measured and the all of the images were normalised to the area of the smallest object in the selection. Images were set to greyscale and the luminance histograms were matched for each image using the SHINE toolbox (Willenbockel et al., 2010). The background colour was set to the average luminance across all images (see **Figure 4.2**).



**Figure 4.2.** Selection of object images used in the adaptation experiment. Images are normalised to the same foreground area and luminance matched. The background was set to the average luminance within each of the images.

#### 4.2.4 Adaptation scan

Eight, luminance matched, size normalised, object images were used to create eight size shift conditions containing four same-different pairs. Each condition was repeated eight times. For each instance of a different condition, all eight objects were used within the block, for every repetition the order of presentation was randomised, with the same order never repeated within the experiment. For each of the eight repetitions of a same objects condition, each of the eight available objects was used. In total throughout the experiment, every object from the

selection will appear exactly the same amount of times in the same and different object conditions, a total of 256 times in each.

Condition 1: 0% area change, different objects

Condition 2: 16.7% area change, different objects

Condition 3: 33.4% area change, different objects

Condition 4: 50.1% area change, different objects

Condition 5: 0% area change, same object

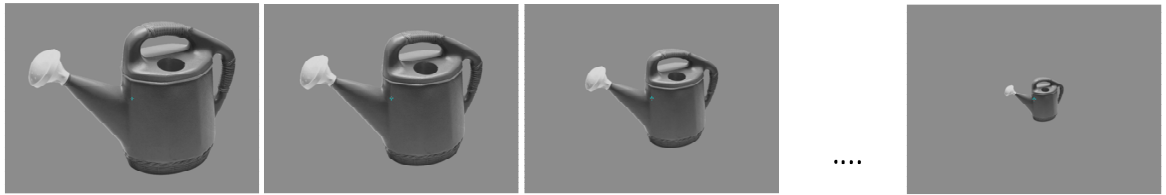
Condition 6: 16.7% area change, same object

Condition 7: 33.4% area change, same object

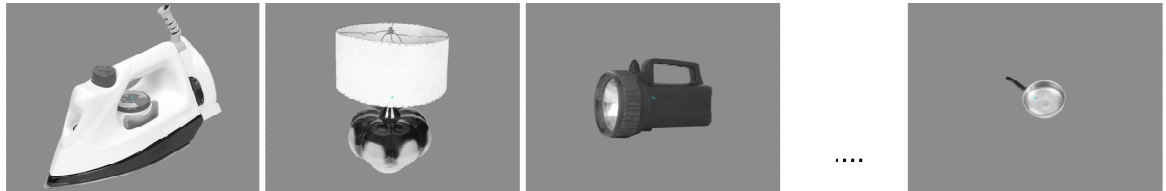
Condition 8: 50.1% area change, same object

Size change manipulations were created by sizing each successive image within the block to one of eight possible sizes ranging from  $3^{\circ 2}$  to  $15^{\circ 2}$  of visual angle. The same eight sizes were used across all size change blocks and as such they have identical average image sizes and visual field stimulation across the block. To achieve varying size changes, the order of the image presentation was manipulated such that the size change between successive images was an average of one of three possibilities, 16.7%, 33.4% or 50.1% area change. This was done by generating unique random permutations of size orders and selecting those with appropriate mean differences between successive image sizes. Examples showing how this is accomplished for both small (17.6%), and large (50.1%), size change conditions can be seen in **Figure 4.3**.

Same object, average size change 16.7% (from  $2^\circ 2$  to  $15^\circ 2$ )



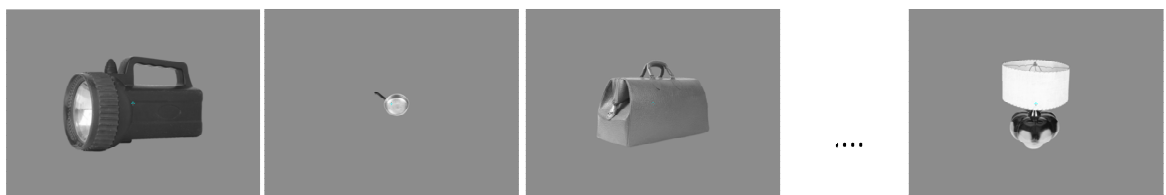
Different objects, average size change 16.7% (from  $2^\circ 2$  to  $15^\circ 2$ )



Same object, average size change 50.1% (from  $2^\circ 2$  to  $15^\circ 2$ )



Different objects, average size change 50.1% (from  $2^\circ 2$  to  $15^\circ 2$ )



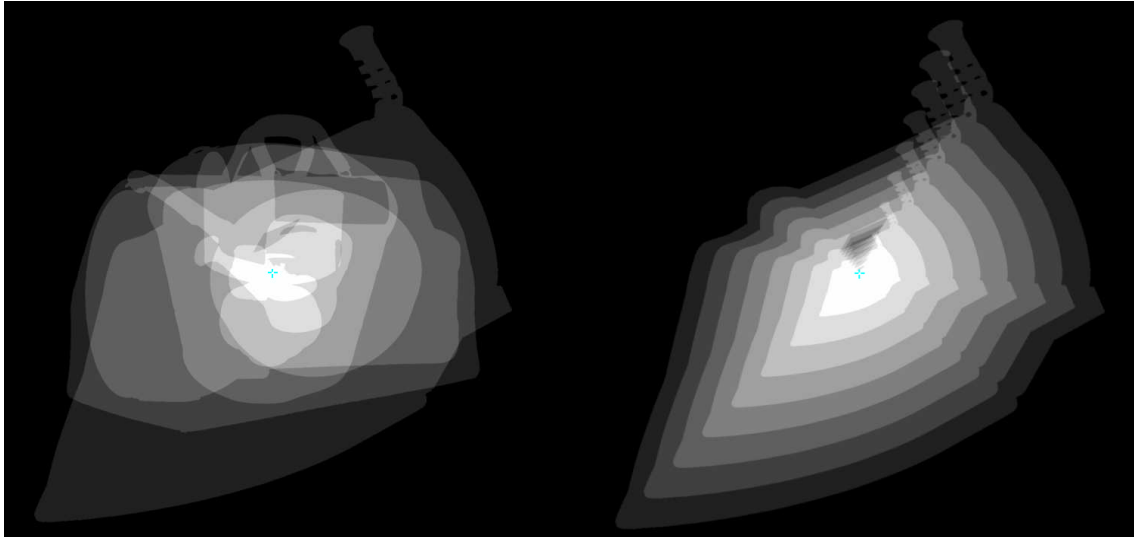
**Figure 4.3.** Low & high size change examples (fixed mean size  $8.5^\circ 2$ ).

Illustrates how different mean size change within a block can be achieved by re-ordering a sequence of object sizes. The top two rows show an example of a 16.7% size shift sequence applied to same and different identity objects. Individual changes are small resulting in a relatively ordered change from big to small. The bottom two rows show an example of a 50.1% size shift sequence applied to same and different identity objects. Individual changes are large resulting in a more disordered sequence of big and small objects. Also included were 0% and 33.4% same and different size change blocks (not shown).

The 0% change condition maintained the same area throughout the block, and these images were sized at the mean area from all other size change blocks,  $10.5^\circ 2$  of visual angle. Because image area was controlled across conditions it was possible to apply the same size change manipulations to both same and different conditions (**Figure 4.4**). Each condition pair fell into one of the four size change conditions from 0%, up to 50.1% and they are linearly spaced apart such that they

represent a parametric manipulation of size change. Within a block, each image was presented for 900ms and separated by 225ms fixation cross in order to reduce the percept of smooth movement. Each block consisted of eight images, lasting nine seconds in total. The blocks were separated by fixation intervals lasting nine seconds.

The nature of the present experiment requires careful control of visual field stimulation, as well as participants' maintained attention throughout the scan. As such a simple attention demanding task was used, which required continuous fixation throughout the stimulus blocks. A fixation cross subtending no more than  $1^{\circ 2}$  of visual angle was displayed on the screen at all times. Pseudo-randomly throughout a block, the inner pixels of the four line segments comprising the cross changed colour to red pseudo-randomly throughout the block. The participants' task was to detect the change in the upper vertical segment of the cross with a button press. Targets were constrained to appear on any but the first image in a block. Each block contained either one or two red line targets, distributed uniformly across conditions, meaning approximately 19% of all images within the experiment contained a red line target. It was found that performance on this task was severely impaired when deviation from fixation by any more than  $3^{\circ}$  of visual angle away from the centre. Results of the attention task were used primarily to verify attention and fixation and were not considered in the main analysis. An accuracy level lower than 80% on the fixation task was deemed to be an appropriate exclusion criterion.



**Figure 4.4.** *Object image silhouettes.*

*(left) shows the range of overlapping visual field stimulation for a size shift block of different objects shown at all possible sizes. (right) shows the range of overlapping visual field stimulation of a size shift block of a single object shown at all possible sizes.*

#### 4.2.5 Localiser scan

Subjects viewed 20 blocks of 10 images. Each block contained images from one of six different categories: faces, bodies, places, objects, scrambled objects, or Fourier scrambled images of the former categories. Face images were taken from the Psychological Image Collection at Stirling (PICS; <http://pics.psych.stir.ac.uk>) and body images were taken from a body image collection at Bangor (Downing et al. 2001; <http://www.bangor.ac.uk/~pss811/page7/page7.html>). Images of other categories were taken from a variety of web-based sources. Each image was presented for 700ms followed by a 200ms fixation cross. Stimulus blocks were separated by a nine second fixation screen. Each condition was repeated four times, and arranged in a counterbalanced block design. In order to maintain attention participants were asked to respond to a red dot superimposed at a random location on a subset of the localiser images, occurring once or twice per block.

#### 4.2.6 *fMRI analysis*

##### *Whole brain*

Whole brain analysis of the fMRI data was carried out using FEAT (<http://www.fmrib.ox.ac.uk/fsl>; Smith et al., 2004). The initial nine seconds of data from each scan were removed to minimize the effects of magnetic saturation. Motion correction procedures were carried out, along with removal of non-brain structures, spatial smoothing (Gaussian, FWHM 6mm) and temporal high-pass filtering (cut off, 0.01 Hz). A two-step mixed-effects analysis was carried out. An initial fixed effects model of each of the eight conditions relative to inter-block fixation baseline was used to analyse individual participant data. Group analysis was carried out using a random-effects model (FLAME, <http://www.fmrib.ox.ac.uk/fsl>). Registration to high-resolution individual space and standard MNI space was done using FLIRT (FMRIB's Linear Image Registration Tool) for each individual (Jenkinson et al., 2002). Contrasts were calculated for each size change condition with same objects compared to their respective size change conditions containing different objects. A linear trend contrast across the four adaptation pairs was calculated. The degree of significance within a voxel for this contrast reflects the extent to which adaptation increases linearly with linearly increasing size change, or size change sensitivity. A contrast of all size stimulus blocks vs. fixation baseline was calculated in order to aggregate activity in all conditions and show regions selective for visual stimuli.

##### *Regions of interest*

The contrasts of objects>scrambled objects from the localiser scan were used to define the LOC region of interest (ROI) in the adaptation scan, and places>faces was used to define the place selective PPA. All functionally localised regions were defined in each individual by anatomically constraining the activation using probability maps from the Harvard-Oxford cortical structural atlas (<http://www.fmrib.ox.ac.uk/fsl/fslview/atlas-descriptions.html>), and selecting up to

the 50 most active voxels within this mask. A minimum threshold of  $P < 0.001$  (uncorrected) was applied, and individuals not showing activation above this threshold for a given region were excluded from the analysis. The V1 ROI was defined anatomically by manually masking the calcarine sulcus of each individual structural scan, and further constraining the mask using active > rest in the localiser scan, ensuring that the selected voxels were responsive to the area of the visual field stimulated throughout the experiment.

GLM parameter estimates for each adaptation scan condition were extracted for each ROI and all voxels within each region were averaged together. Activation values for each region were converted from arbitrary units to units of % signal change relative to fixation baseline ( $\text{parameter estimate} / 100 * \text{Regressor Height}$ ). Responses from all ROI's were subjected to repeated measures ANOVA to determine significant differences across regions and identity.

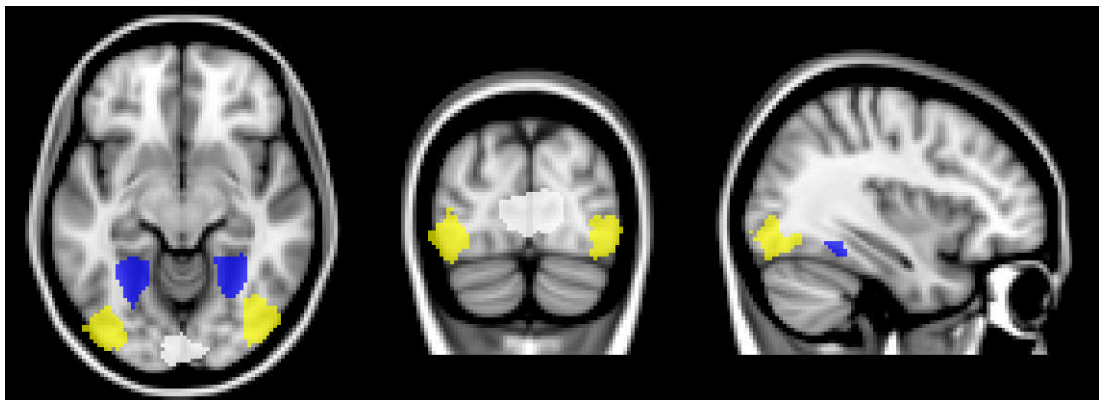


### 4.3 Results

#### 4.3.1 Localiser

*Object and place selective regions of interest were identified within the (VOT) region of the brain (Figure 4.5;*

**Table 4.1).** The LOC was identified as a region in the lateral occipital cortex or posterior cingulate showing a greater selectivity for objects relative to scrambled objects (Malach et al., 1995). The PPA was identified as a region along the parahippocampal gyrus that showed a greater selectivity for place stimuli relative to face stimuli (Epstein & Kanwisher, 1998). Both place and object contrasts elicited sizable bilateral regions of activation in all participants. The V1 ROI was created anatomically by manually identifying the calcarine sulcus in each individual and constraining the region using a functional contrast of active>rest in the localiser scan. This ensured that only active portions of the visual field were sampled. Masks were created for each of the above regions which were then used to extract mean time series' for each region within the adaptation scan for all six conditions.



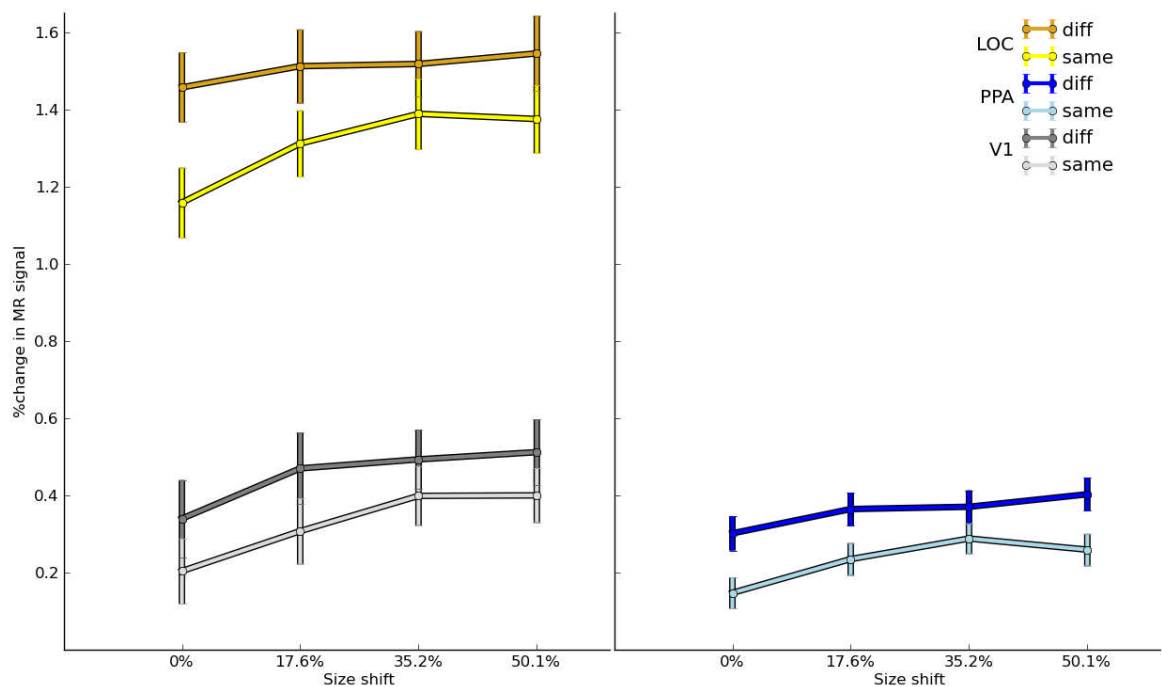
**Figure 4.5.** *Category selective, localiser scan ROI masks shown in MNI space. Yellow=LOC (objects>scrambled), blue=PPA (places>faces), white=V1 (anatomical calcarine sulcus combine with active>rest). Masks are created using a minimum threshold of  $P<0.001$  (uncorrected). Images are centred on MNI coordinates: 36, -82, -10. Scans are displayed using neurological convention with the right hemisphere displayed on the right.*

**Table 4.1.** *MNI coordinates of regions of interest peaks*

Area	Hemisphere	MNI Co-ordinates			Z-value
		X	Y	Z	
LOC	Right	38	-80	-18	7.27
	Left	-40	-84	-12	7.26
PPA	Right	22	-46	-16	7.68
	Left	-24	-56	-14	7.48
V1	Right	10	-82	-6	7.88
	Left	2	-84	-8	7.36

#### 4.3.2 *Size sensitivity in object selective cortex*

All participants included in the analysis performed the attention demanding central fixation task well within acceptable levels of error (Mean target detection rate= 90.3%, SD=6.6%) and as such were assumed to have maintained stable fixation and sustained attention throughout the adaptation scan. One individuals' data were excluded from the analysis due to excessive movement and poor performance on the fixation task.



**Figure 4.6** BOLD responses for same and different identity objects in all ROI's. Response profile shown for LOC (yellow), V1 (grey), and PPA (blue). Measures of %signal change relative to fixation baseline were averaged across hemispheres and participants. Error bars represent  $\pm$  standard error of the mean.

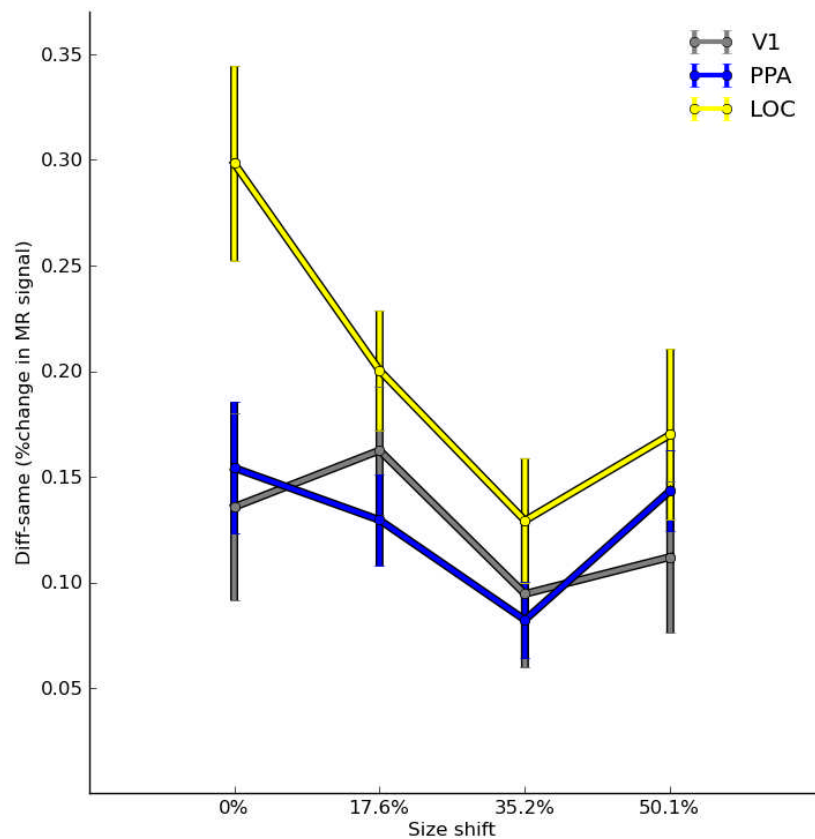
No interactions involving hemisphere were found in any ROI, there was a magnitude difference in the response of the PPA, with the response higher in the right than the left hemisphere. The pattern of response across conditions in the PPA was the same across hemispheres. No main effects of hemisphere were found in the LOC or V1. Responses across hemispheres were pooled for all subsequent analysis.

The responses for all size shift conditions for both same and different identity objects for all ROI's can be seen in **Figure 4.6**. A 3x2x4 ANOVA (region, object identity, size shift) was carried out. There was a main effect of region ( $F^{(2,50)} = 86.907$ ,  $P < 0.001$ ), a main effect of object identity ( $F^{(1,25)} = 78.082$ ,  $P < 0.001$ ), and a main effect of size shift ( $F^{(3,75)} = 17.518$ ,  $P < 0.001$ ). There was an interaction of region with object identity ( $F^{(2,50)} = 10.353$ ,  $P < 0.001$ ), region with size shift ( $F^{(6,150)} = 2.215$ ,  $P < 0.05$ ), but no interaction of object identity by size shift ( $F^{(3,75)} = 2.375$ ,  $P = 0.78$ ). The three way interaction of region by object identity by size shift was

significant ( $F^{(6,150)} = 2.702, P < 0.05$ ). The results suggest that overall the response varies by region, same identity objects result in a lower response relative to different across all regions, and the response in all regions is modulated by size shift. The effect of object identity interacts with regions, suggesting, that in one or more of the regions, the same identity adaptation is greater than the others. The identity, size shift interaction suggests that over all regions, the effect of size shift is greater for same identity objects than for different. The three way interaction indicates that the pattern of size shift modulation across identity conditions varies across regions.

In order to examine specific effects within each region separately, a linear trend statistic was computed across the four size shift conditions. All three regions showed significant main effects of linear trend (LOC:  $F^{(1,25)} = 34.440, P < 0.001$ , PPA:  $F^{(1,25)} = 28.964, P < 0.001$ , V1:  $F^{(1,25)} = 19.939, P < 0.001$ ). Only for the LOC was there an interaction of linear trend with identity (LOC:  $F^{(1,25)} = 14.041, P < 0.001$ , PPA:  $F^{(1,25)} = 0.434, P = 0.516$ , V1:  $F^{(1,25)} = 0.736, P = 0.399$ ). These results further suggest that in general, the response shows a linear increase across size shifts. However, in the LOC, the response to same identity objects shows a stronger linear increase in response with increasing size shift compared to the response to different identity objects. Both non-object selective regions show an equivalent linear trend regardless of object identity.

The above results are summarised in **Figure 4.7**, whereby adaptation responses were computed across the four size shift conditions by subtracting the response to same identity objects from the response to different identity objects. A linear trend contrast of adaptation across the four view shift conditions was significant only for the LOC ( $F^{(1,25)} = 12.958, P < 0.001$ ). Adaptation was high for the low size shift conditions and decreased roughly linearly through higher size shift conditions. This was not true for either of the non-object selective regions.

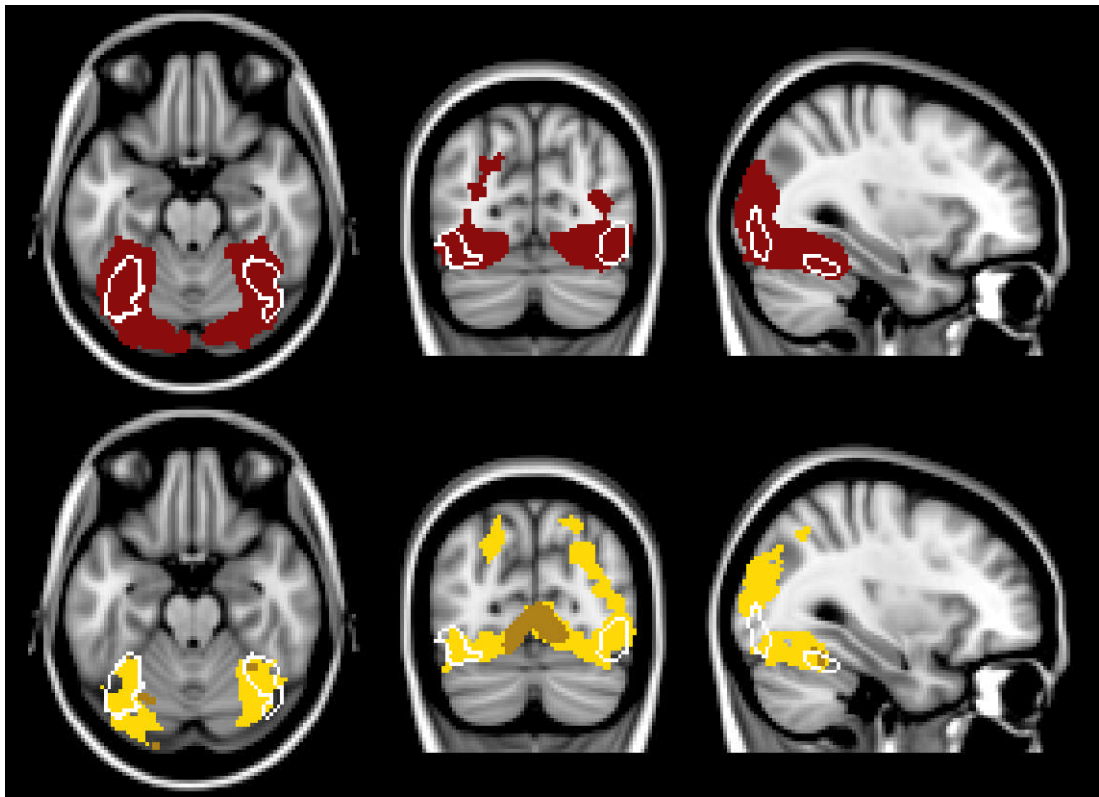


**Figure 4.7.** Adaptation across size shift conditions in all three ROIs. Adaptation computed as the response to same identity objects subtracted the response to different identity objects. Data shown for LOC (yellow), V1 (grey), and PPA (blue). Measures of % signal change relative to fixation baseline were averaged across hemispheres and participants. Error bars represent  $\pm$  standard error of the mean.

#### 4.3.3 Whole brain size sensitivity

**Figure 4.8** shows both adaptation for same objects relative to different, regardless of size shift, and linear trends across the four size shift conditions for both different and same identity objects. These comparisons are shown in the whole brain and include the borders of the functionally defined LOC. Considerable same identity adaptation is seen throughout the early visual and category selective cortex, and while a linear trend is extensive for same identity objects, it is restricted to medial and posterior regions of the VOT for different identity objects. These results are largely in agreement with pattern of responses seen in the ROI results; however, one observation that can be made is that there are spatial differences in size

sensitivity within the LOC. There are small bilateral regions in the most anterior portion of the LOC showing additional size sensitivity for different objects.



**Figure 4.8.** Whole brain adaptation and linear trends.

(Upper slices) Areas in red are voxels showing adaptation to same objects relative to different objects regardless of size shift. (Lower slices) Voxels showing a significant linear increase in response with increasing size shift for both different identity objects (dark orange), and same identity objects (light orange). Outlines of the LOC as defined by a contrast of objects > scrambled are shown in white. Statistical images are created using a threshold of  $P < 0.001$  (uncorrected). Images are displayed on the MNI152 standard brain, centred on MNI coordinates: 32, -74, -22. Scans are displayed using neurological convention with the right hemisphere displayed on the right.

#### **4.4 Discussion**

The present study used a parametric adaptation design to investigate the effects of varying stimulus size and identity, on adaptation responses in object selective cortex, while holding total visual field stimulation constant. Comparing non-object selective control regions, retinotopic primary visual cortex, and place selective PPA, with the object selective LOC region, it was found that all regions showed neural adaptation when object identity was held constant relative to when identity was varied. The responses in all regions were modulated by size shift. However, only in the LOC was the pattern of response across size shift also modulated by object identity.

These results suggest that it is possible to show same-object size-invariant adaptation in object selective regions, even at large size shifts. The amount of adaptation however, is modulated by the degree of size change. The representation of objects within the LOC remains sensitive to the degree of size change for objects in the visual field as indicated by linear trends across size shift conditions largely restricted to same identity blocks. The pattern of response in the LOC is not consistent with a size invariant representation of objects, and in fact closely matches the hypothesized profile of a size-dependent region (**Figure 4.1**). The present finding is consistent with the monkey single unit literature in which a greater proportion of cells in object selective cortex are found to be sensitive to size change than invariant to it (Miyashita & Chang, 1988; Schwartz et al., 1983).

Apparent same identity adaptation was found in non-object selective regions, including PPA and V1, however, this was thought to reflect adaptation due to low level visual features, which vary less for same identity blocks regardless of shift, was found in non-object selective regions including V1 and PPA. The modulation of the adaptation response in the LOC was found to be over and above that driven by low level visual features (indicated by interactions between size shift, identity and region).

*4.4.1 LOC as an intermediate stage in object representation*

The present experiment independently manipulated both object identity and size shift while controlling for the extent of visual field stimulation across conditions. The response profile in the LOC showed an interaction between these two properties. The modulation of the LOC by both object identity and size shift is consistent with the idea that the representation in the LOC is an intermediate stage in establishing an object centred coordinate frame. In an object centred coordinate frame the features of an object are coded relative to important features of the object, rather than to the viewer. Establishing such a representation scheme has been suggested to be necessary in order to perform efficient object recognition (Marr & Nishihara, 1978; Marr, 1983). In such a scheme, a given neuron may code for a particular feature of an object, and if that feature is present then the neuron would fire equally regardless of where in the visual field that feature occurred. Such a representation would be invariant to changes in size, view, or position (Biederman & Gerhardstein, 1993) .

A general principle thought to underlie coordinate transformation in the brain is gain modulation (Salinas & Thier, 2000). This is a principle initially described in posterior parietal neurons (PPC) that were found to respond to a particular point in the visual field, but their response magnitude, or gain, was also modulated by the position of the eyes (Andersen & Mountcastle, 1983). Gain modulation is defined as the change in the amplitude of a particular neurons' response without a change in its selectivity or receptive field characteristics. The source of this modulation is input from a different sensory or cognitive modality and it allows multiple sources of information to be combined and simultaneously represented in a single neuronal population (Abbott, Chance, & Salinas, 2008).

The utility of these cells in performing coordinate transformations was first described by Zipser and Andersen (1988) who trained a neural network simulation to perform a transformation of retinal coordinate frame, into a body centred coordinate frame, given the input of retinal position and gaze angle. When the



network learned to correctly perform this transformation, the units in the hidden layer were examined, and found to be similar to the gain modulated receptive fields found in the PPC. Further confirmation of the likely contribution of gain field representations was provided by Pouget and Sejnowski (1997), who showed that lesioning of a model of PPC gain field function produced similar visual and spatial deficits to those found in patients with posterior parietal lesions, in particular, deficits effecting multiple coordinate frames.

Gain fields have been described in many other regions outside of the parietal cortex. In particular they have been described in ventral visual regions, and have been suggested to underlie various kinds of coordinate transform useful for invariant object recognition. Connor et al., (1996; 1997) recorded from shape selective V4 neurons in monkeys trained to direct their attention to specified parts of the visual field. They found that the responses were retinotopically organised and position sensitive, but were gain modulated by directed attention, suggesting that this population can be used to compute a representation in attention centred coordinates that is invariant to position in the visual field (Salinas & Abbott, 1997).

In addition to attentional gain modulation, Dobbins et al. (1998) recorded from Macaque V4 neurons while presenting object stimuli at different distances and different physical sizes. They found that neurons in this region are sensitive to particular retinal sizes of objects, however, their responses were found to be gain-modulated by viewing distance, as determined by fixation distance and binocular disparity information. This gain modulation could be an additional source of information used to compute a size invariant representation later in the visual hierarchy, indeed, lesion experiments of this region suggest that it is crucial for making size judgements (Humphrey & Weiskrantz, 1969; Ungerleider, Ganz, & Pribram, 1977).

The present experiment contained a demanding central task in which participant attention remained narrowly focussed on the central fixation cross throughout all conditions. The stimuli were presented on a back projection screen through a mirror in the scanner, ensuring fixation distance was also held constant

for all participants. Neural adaptation in the LOC is driven by object selectivity, and shows an attenuation of signal when identity is held constant, however, there is a modulation of the adaptation response in the LOC whereby greater shifts in size result in a reduction in adaptation. In the absence of attentional modulation and binocular depth cues, both V4 and LOC appear to remain sensitive to size change. The present results are consistent with such a gain modulated, intermediate representation, allowing a size invariant representation to be computed by downstream regions. The nature of the relationship between fMR-adaptation and gain modulation population responses have yet to be described, therefore this interpretation remains speculative. Nonetheless, it is possible to predict that without an attention demanding central task, and with participants explicitly instructed to pay attention to the object regardless of size, the response in LOC across size shift conditions would show smaller modulation by size shift.

#### *4.4.2 Adaptation in retinotopic cortex*

In all conditions in the present experiment, the average visual field stimulation across any given block was held constant. Nonetheless it was found that all regions, including non- object selective primary visual cortex and PPA, showed neural adaptation when object identity was held constant relative to when identity was varied.

The simplest interpretation for the considerable adaptation seen throughout the visual system when comparing these conditions is that smaller amounts of visual change in the same object condition led to image level adaptation in all visually responsive regions. The images used in the same and different conditions were luminance matched, and the same object images were used to construct all sets of conditions. This was to ensure that these conditions were well matched in terms of low level visual features. The amount of visual change between successive images within the same and different blocks was however, unavoidably greater for the different condition relative to the same condition (**Appendix C. 3**). The pattern

of adaptation seen in LOC for same identity objects was found to be greater on average than in non-object selective regions, and was modulated by size shift. It may be possible to devise stimulus manipulations that adjust the amount of low level visual change in same identity conditions, such that they match the change in different identity conditions. However, this would involve introducing visual noise, which could create additional confounds of recognisability. Few studies of invariance in category selective regions consider the image level contribution to fMR-adaptation. There is evidence that a proportion of category selective adaptation is driven by low level feature properties, that differs depending upon the visual category in question (Andrews et al., 2010). The present study, with multiple size shift conditions, comparable in terms of visual field stimulation, is able to show image level adaptation across all retinotopically organised regions, and adaptation over and above this in object selective LOC; highlighting the need to consider the contribution of image level visual properties to adaptation responses.

A further consideration for the interpretation of adaption found in primary visual cortex concerns the relative contribution of foveal and peripheral voxels to the overall ROI response. The V1 ROI contains a full retinotopic map of the visual field, and voxels corresponding foveal and peripheral eccentricities would be expected to respond differently to varying sizes of stimuli, with foveal voxels being consistently active and peripheral voxels only receiving stimulation at all when large object images are in view. In the different objects condition the intensity of stimulation in the foveal visual field would change dramatically as each object image is presented. In the same objects conditions there would be less change in intensity of stimulation within foveal voxels due to the central part of an object image not changing substantially as the image is resized, and this could result in an attenuation of response consistent with neural adaptation. In a conventional ROI analysis all voxels from both foveal and peripheral eccentricities are aggregated and averaged together. A difference between different and same object conditions might be expected in this situation driven largely by the response in foveal neurons, which are generally more numerous in V1 than peripheral ones. Additionally, large

scale eccentricity biases have been reported across the VOT, with medial regions (including the parahippocampal cortex/PPA) showing a peripheral bias, and lateral regions (including the LOC) showing a foveal bias (Levy et al., 2001; Malach, Levy, & Hasson, 2002).

The present study was designed to take into account and control for the effects of size change manipulations within retinotopically organised visual regions, hence the selection of primary visual cortex as a control ROI. If indeed the pattern of adaptation found in V1 is driven by the limited amount of visual change in the foveal part of the visual field for the same objects condition, it is therefore possible, when examining a single size shift condition, to show an apparently size-invariant pattern of adaptation even in a low level retinotopically organised visual region that shows no selectivity for object stimuli. This effect, coupled with the recent discovery of retinotopic maps in object selective visual regions (Larsson & Heeger, 2006), could account for a great many findings showing size invariant adaptation in these regions. Typically, experiments investigating the issue of size invariant adaptation exert considerably less control over the low level visual properties of their same and different identity object stimuli than the present experiment. Future studies using retinal position to control for retinotopic confounds and a similar design to the present one would be useful in detecting entirely identity based neural adaptation.

#### *4.4.3 Summary and conclusions*

When object identity is held constant the degree of size sensitivity is considerably greater and more extensive than when object identity is varied. This is true within object selective cortex and regions extending beyond it. The monkey literature shows that there are both size invariant and size dependent neurons within IT (Miyashita & Chang, 1988; Schwartz et al., 1983). The present experiment, by considering multiple size shift conditions within a single experiment, provides strong evidence that the population representation of objects at the level of the VOT/LOC is size sensitive. Furthermore, adaptation in the LOC is driven by the high

level object features and geometry and cannot be explained by low level visual change or effects of the extent of visual field stimulation.

The present design measured the response for parametrically manipulated size shifts. Size sensitivity was measured at the population level by examining linear trends across the different levels of the manipulation. Testing for linear trends across shifts can only reveal sensitivity; invariant responses would be flat, and invisible to such a measure. It is possible to further exploit the parametric manipulations used here, and to extend these findings by examining responses at the level of an individual voxel. The slope of a fit across size shift conditions is a measure of a given voxels sensitivity/invariance. This can be combined with independent measures of selectivity and used to visualise at the whole brain level spatial patterns of invariance. Single unit studies suggest a posterior anterior gradient of increasing size invariance (Ito et al., 1995; Lueschow et al., 1994), but due to the limited sampling inherent in such techniques, it is difficult to fully describe this within a single experiment. fMR-a studies suggest a posterior/anterior division within LOC of size sensitivity/invariance (Grill-Spector et al., 1999; Sawamura et al., 2005), but regions of interest methods and univariate analyses are unable to reveal graded change in function. Chapter 5 (section 5.4) unpacks the size sensitivity measures from the present experiment, and explores this visualisation method for same and different identity size shift manipulations, as well as a combined measure of size change adaptation.

## **5 Mapping invariance in category selective cortex**

### **5.1 Introduction**

The aim of the present thesis is to investigate how the process of establishing invariance influences the organisation of the VOT region. The previous experimental chapters have developed a paradigm whereby visual change is parametrically manipulated across multiple conditions in order to measure the extent to which a given region is invariant to such changes. The parametric adaptation paradigm has been used to perform comparisons of the pattern of response between different functionally defined category selective cortical regions, and for different visual categories. It has been possible to show how the degree of invariance varies depending upon functional properties of the cortical region, and the visual category being viewed.

An advantage of the parametric adaptation paradigm is that it can be used not only to look at the response in a functionally defined region, but also at the level of the individual voxel. The brain response in category selective voxels, across multiple parametrically related conditions, reflects the extent to which those voxels are invariant, or sensitive, to a particular stimulus property. The present chapter aims to characterise this response in order to map how invariance changes over space within category selective cortex. This measure of invariance is comparable across the three experiments described so far (chapters 2, 3, & 4), it is therefore possible to extend the findings of previous chapters, and show how maps of invariance change depending on stimulus category, and the nature of change in the visual input.

5.1.1 *Mapping without statistical thresholds*

The term ‘mapping’ in the context of neuroimaging implies the localisation of particular brain functions. However, according to a review by Jernigan, Gamst, Fennema-Notestine, and Ostergard (2003), most functional neuroimaging studies do not apply the appropriate statistical comparisons in order to make inferences about the spatial arrangement of functional differences. It is common practice in neuroimaging experiments, to present whole brain maps of ‘brain activations’ with statistical thresholds applied. Such activation maps are overlaid in colour on a structural image displaying the anatomy of the brain. Islands of statistically significant activity are often discussed in terms of the active region being significantly activated by a given task, and the functional properties of that region are then inferred. Jernigan et al., (2003) argue that regions adjacent to significant activations, which may be below a conventional statistical threshold, would in fact not be significantly different from the activated region. What is needed in order for such inferences to be valid, is a comparison of the region in question with the rest of the brain. Given the aim of brain mapping, to localise function, it is necessary to show both regions that are, and *are not* active.

The use of common statistical thresholds applied by convention across studies is designed to strike a balance between Type I and Type II error. Whilst this convention may be useful when considering the outcomes of many studies, in individual cases it can lead to spurious reasoning, namely, that individual findings that narrowly fail to reach significant are due entirely to chance, and should be ignored. The use of such a thresholding procedures, pioneered by Fisher (1956), were not intended to be applied in this way. “...in fact no scientific worker has a fixed level of significance at which from year to year, and in all circumstances, he rejects hypotheses; he rather gives his mind to each particular case in the light of his evidence and his ideas...A test of significance contains no criterion for ‘accepting’ a hypothesis” (Fisher, 1956).

Lieberman and Cunningham (2009) suggest that the current trend towards increasingly conservative statistical thresholds in neuroimaging research (in order to reduce false positives) leads to an increased Type II error rate and an increase in several kinds of bias in neuroscientific findings. The high costs of fMRI research and the typical sample sizes used means that the statistical power available for many researchers restricts findings to large and obvious effects. This leads to a bias towards motor and sensory investigations and to the exclusion of investigations into more subtle brain processes. A further problem that can arise is that effect size measures are only reported when activations are significant. If this happens less often because of increased stringency, then it is less likely that meta-analytic techniques will be able to aggregate across multiple studies and uncover subtle but meaningful effects. Conversely, meta-analytic techniques would ultimately reveal any false positives for what they are through aggregation. It is argued therefore, that the nature of the threshold applied should be tailored to the type of desired inferences arising from a particular study. In certain cases it may be particularly important to prioritise the reduction of Type I error, for example, when conducting intervention studies to choose between particular drugs or treatments. However, when conducting basic exploratory research the focus should be on replication and meta-analysis, whereby Type I errors are eliminated by their failure to replicate (Lieberman & Cunningham, 2009).

The present chapter develops methods of whole brain visualisation which are meaningful without any statistical thresholding. Such continuous unthresholded maps would not, in themselves, allow statistical inference based on direct comparisons of spatially distinct brain areas. Therefore, it is also necessary to devise methods that can test for the reliability of spatial patterns, and allow valid conclusions to be drawn. The measures of invariance that can be devised from previous chapter's data can be assumed to relate directly to an abstract property of the stimulus. The main aims of this chapter are: - 1.) To quantify invariance at the level of the individual voxel in order to generate continuous, unthresholded whole brain maps. 2.) to devise statistical approaches that can be used to assess the



reliability of spatial patterns of invariance and selectivity throughout the ventral visual processing hierarchy. 3.) to show that unthresholded but statistically robust whole brain maps can reveal previously invisible patterns within the VOT.

### *5.1.2 Selectivity and Invariance in the VOT*

As discussed in the introduction chapter, it was once thought that the VOT was organised primarily on the basis of category selectivity, with islands of activity corresponding to functionally distinct cortical modules (Kanwisher, 2000). More recent findings however suggest that there is a continued and coarse scale influence of retinotopy still present at the level of the VOT (Levy et al., 2001). Furthermore there is evidence that representational distinctions exist at finer spatial scales that cannot be discerned using univariate techniques. Multivariate techniques, such as those pioneered by Haxby et al. (2001) are able to distinguish between stimulus categories from activation in the VOT region. This was true even excluding category selective cortical regions from the analysis. Recently, Haxby et al. (2011) extended these multivariate techniques and described a new ‘hyperalignment’ method in which brain response patterns from viewing a full-length movie, are used to transform individual brain activation data into a common, high dimensional space corresponding to the representation space of complex visual stimuli. Haxby et al.’s approach is able to achieve excellent classification performance between subjects using this new method, where previously only effective within-subject classification was possible. This suggests that the model is a useful approximation of the representational space used by the brain to encode visual stimuli. However, it remains difficult to understand exactly what individual dimensions within such a putative high-dimensional representation space might be. The parametric adaptation paradigm developed throughout this thesis attempts to devise ways to sample within such a space by manipulating high level properties of the visual stimulus.

Perception and recognition are performed by neurons in a hierarchical network progressing from the purely visually responsive V1, through to the medial temporal lobes, implicated in memory processes (e.g., Squire, Stark, & Clark, 2004). At each successive stage of the visual processing hierarchy, the receptive field size of a neuron at that stage increases from responding to a small patch of the visual field to the point where appropriate stimuli anywhere in the visual field will drive the response (Rolls, 1994). Cells become increasingly invariant to various kinds of irrelevant change in the stimulus whilst simultaneously becoming increasingly selective for more and more complex and abstract feature conjunctions (Kobatake & Tanaka, 1994; Rolls, 1994). According to Ison & Quiroga (2008), the basic principles of neural coding in the ventral visual pathway rely on the orthogonal properties of *Selectivity* and *Invariance* increasing at each successive stage, moving gradually from retinotopically organised representations of the input, to abstract representations that facilitate recognition and memory.

Many neuroscientific techniques are designed to measure neuronal selectivity of one form or another, and techniques such as fMR-adaptation among others can be used to measure invariance. Measures of selectivity and invariance, and how they change over space, have not previously been combined into a single analysis, and the present chapter aims to do exactly this. For each of the experiments conducted throughout this thesis, measurements have been taken that can be used to derive metrics of both selectivity and invariance. Each experiment included an independent functional localiser scan, designed to define category selective regions of interest. At the whole brain level, a contrast of any given category condition with a scrambled counterpart, constitutes a measure of selectivity for that category for each voxel. Furthermore, in the parametric adaptation scan for each experiment, the linear trend response of each voxel across all four shift conditions, describes the extent to which what voxel is sensitive to change in the stimuli. It is therefore possible to compute a metric of invariance for each voxel in the brain. These two measures are combined in such a way that the voxels' colour is determined by the degree of invariance and the opacity is

determined by the degree of category selectivity. This allows for the creation of whole brain maps that are capable of showing graded change in both selectivity and invariance without any statistical thresholding.

### 5.1.3 *Spatial comparisons of invariance*

The second aim of the chapter is to devise statistical means to show that invariance changes at different levels of the visual processing hierarchy. Conventional whole brain statistics do not compare the activity in one brain region to that of another spatially distinct region. Typically whole brain statistics are used to perform independent comparisons between conditions for each voxel and show those voxels which reliably show greater activity in one condition of interest relative to some baseline condition. Attempts to perform spatial comparisons between brain regions are uncommon (Jernigan et al., 2003). One notable exception is Spiridon, Fischl, and Kanwisher (2006), who investigated the spatial profile of category specific regions within the human VOT. They devised a new method of spatial analysis to determine the extent to which selectivity of each category selective region drops off outside of their functionally defined borders. The spatial analysis involved projecting the outline of the category selective regions onto an inflated cortical surface then selecting voxels corresponding to an outline of the region, 1mm away from the outer edge, and 1mm inside the edge. This process was repeated to generate additional concentric rings extending from 2mm inside the edge to 14mm outside. Rings were generated for each individual and the percent signal change for the voxels in each ring in each condition was averaged across subjects. This analysis allowed selectivity to be analysed as a function of distance from the borders of functionally defined regions. It was found that category selectivity in many of the studied regions drops off sharply from the typically defined borders.

The present analysis chapter will use related methods to sample points in space to follow the posterior-anterior progression of the ventral visual processing hierarchy as well as points spanning medial-lateral axis of the VOT region. It will be

## Chapter 5

possible to test for differences in the spatial profile of invariance across all of the experiments conducted throughout this thesis. The profile is expected to vary depending upon the stimulus type viewed and the type of visual change manipulation (view-, or size- shift).

## **5.2 Methods**

### *5.2.1 Participants and Data acquisition*

The present chapter uses data acquired for chapters 2, 3, & 4 and carries out additional analyses across those studies. The methods for data acquisition and pre-processing are described in detail in chapters 2, 3, & 4. FMRI scans in all studies were acquired using the same scanner and stimulus delivery equipment using equivalent imaging parameters. All pre-processing was done as per the standard FSL analysis pipeline (<http://www.fmrib.ox.ac.uk/fsl>). Custom data visualisation tools used in the present chapter were created using modified components of the SPM5 toolbox ([www.fil.ion.ucl.ac.uk/spm](http://www.fil.ion.ucl.ac.uk/spm)).

Participants were drawn in the main from the University of York postgraduate population, and as such, there was some overlap of participation between studies, four participants took part in all three experiments (see Appendix D for individual within subjects analysis across all experiments). In total there were 56 unique participants.

### *5.2.2 Experimental design and stimuli*

All of the experiments conducted within this thesis used a similar adaptation paradigm, in which a particular stimulus property was manipulated parametrically. In the case of experiments 2 & 3, there was parametric manipulation of view shift consisting of four conditions, separated by 5° steps, from 0° to 15°. Essentially the same manipulation was therefore applied to both objects and landscapes shifting in view. Experiment 4 was a modification of this paradigm to investigate size invariance processes in category selective cortex. Here the average size shift within an image block was manipulated and the average size change increased across four conditions in steps of 16.7%, from 0% to 50.1%. With manipulations of image size it was now possible to perform equivalent manipulations to both same identity and

different identity objects. For detailed descriptions of stimulus construction and experimental design see the relevant method sections in chapters 2, 3 & 4. The present analysis makes the assumption that the parametric manipulation of change paradigm can be used to derive comparable measures of both size- and view-invariance.

### 5.2.3 *Measures of invariance and selectivity*

In each experiment, and for every voxel in the brain, a measure of that voxels response relative to a fixation baseline, for all four parametrically increasing shift conditions, has been acquired. Previous whole brain analyses looked for voxels which showed a significant linear trend across these four conditions. It is possible to use this measure (analogous to the slope of a linear fit) to derive a single metric of invariance to the manipulated stimulus property. The slope parameter can be positive or negative, when it is high and positive it indicates a strong positive increase across the four conditions, when it is close to zero it indicates that the voxel in question does not differentiate between stimulus conditions. A negative slope would not be predicted in any voxel that shows category selectivity. The present analysis considers the slope parameter, normalised to fall between minus one and one (relative to absolute max value in the brain;  $x^i/\max(\text{abs}[x])$ ), to be a measure of invariance, where a value of zero indicates an invariant response, and a value of one indicates maximum sensitivity.

Each of the previous experiments included an independent functional localiser scan. The same conditions and stimuli were used across all three experiments. For each experiment it was possible to define equivalent category selective regions of interest. The present analysis uses the localiser data to define independent measures of category selectivity. Typically, selective voxels are localised by the comparison of a particular stimulus type with an appropriate baseline, and voxels which show significant activation relative to this baseline are considered selective for the stimuli. Here the measure of effect size for the

comparison of intact category images with scrambled images is treated as a measure of category selectivity. The effect size statistic is normalised between zero and one, where zero indicates no selectivity, and one indicates maximum selectivity within the brain ( $[(x^i - \min(x)) / (\max(x) - \min(x))]$ ).

#### 5.2.4 *Mapping visualisation techniques*

Conventional neuroimaging analyses display the voxels in the brain that show a significantly higher response in a particular condition relative to some baseline. Such data are presented as whole brain activation maps and attempts are made to infer the functional properties of activated regions. The size and extent of these activations are determined by many factors extrinsic to brain function, such as the arbitrary statistical threshold used to display the data, the signal to noise ratio (which is affected by number of condition repetitions or number of participants in the experiments), and any kind of correction for multiple comparisons applied to the data. The present chapter utilises a unique mapping approach which is designed to map and visualise aspects of brain function in terms of effect sizes. This approach has the potential to reveal graded change in function within and between classical category selective visual regions that would be obscured in classical thresholded maps of statistical parameters. Conventional statistical thresholding approaches have proven useful for forming valid inferences, but under some circumstances it can disguise underlying spatial structure present in effect size measures. For illustrative purposes the maps generated using the current techniques include visualisations of the borders of classically defined category selective regions, generated at the group level applying statistical thresholds. This allows the relationship of effect size measures and functionally defined regions to be considered.

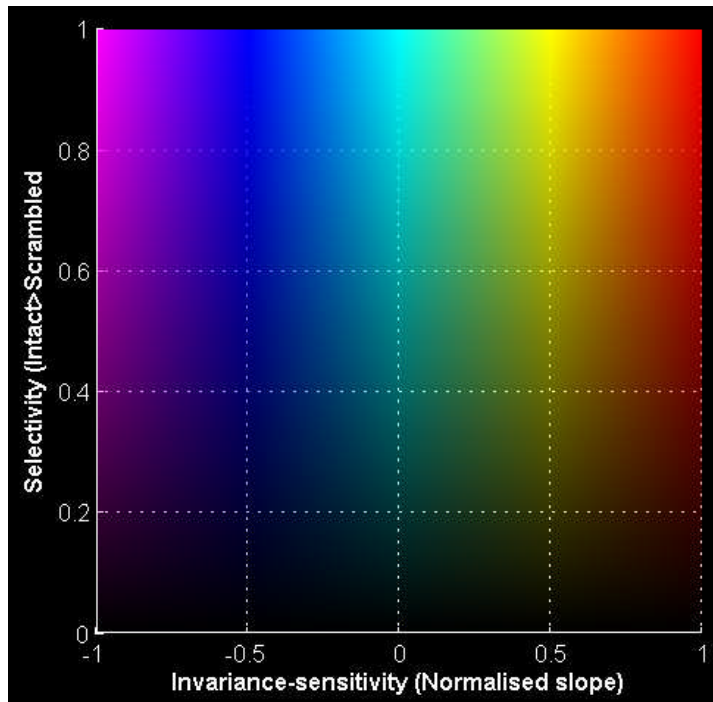
Measures of invariance and selectivity were defined, which are expected to vary and interact in complex ways throughout the ventral visual system. These measures are available for every voxel in the brain, for each of the three

experiments conducted. The best way to visualise these measures in a way that can be meaningfully interpreted is through the use of maps of colour and opacity. The value of one measure, such as invariance, can be used to determine the voxels position along a colour continuum. The level of transparency can then be determined by the value of a second variable. The resulting whole brain map would show graded change in both variables as well as graded change in how they interact. There are many questions that can be asked of the available data, and many ways in which the different measures from each of the experiments can be combined to produce meaningful colour maps. The present analysis focuses on two complementary approaches for determining the colour and opacity of voxels in the brain, invariance x selectivity colour maps, and the vector colour mapping approach.

#### *5.2.5 Invariance x selectivity colour maps*

This method uses the value of a variable of interest to select a colour from a unidimensional colour spectrum created by sampling from the 'HSV' colour map. The custom colour map varies hue systematically, starting from magenta, passing through blue, cyan, yellow, and ending on red. A second, orthogonal, variable of interest is used to determine the opacity of the voxel. This method is useful to examine a particular variable of interest, such as invariance, within a single experiment. The second variable of interest, in this case selectivity, is in units that are not directly comparable, and is used to determine voxel transparency (**Figure 5.1**). This method is used here to visualise in detail how invariance changes within voxels selective for a single visual category.





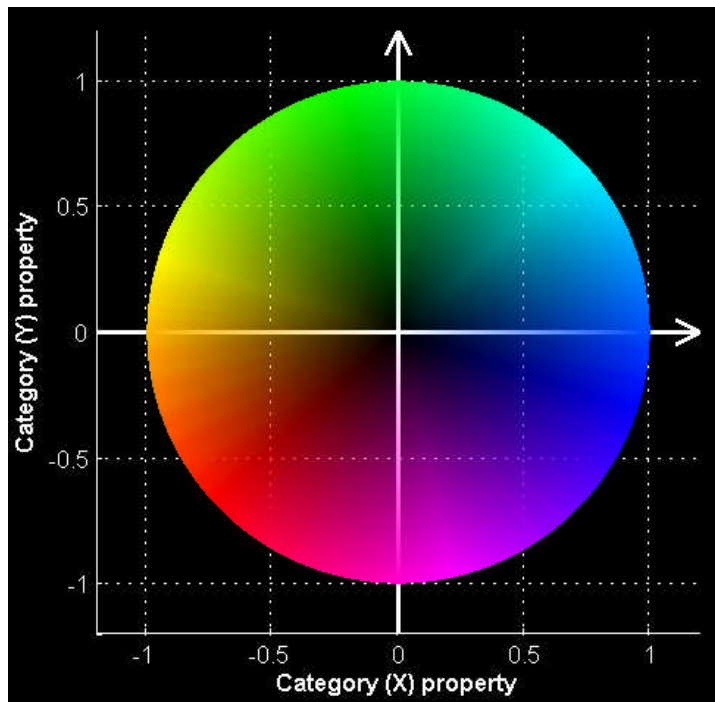
**Figure 5.1.** *Invariance x selectivity colour map.*

*Colour in the sampled HSV colour scale determined by normalised slope parameter (-1 to 1) of a linear function fitted across shift conditions. Opacity determined by normalised effect size (0 to 1) of intact>scrambled comparison in functional localiser.*

### 5.2.6 Vector colour mapping

This mapping approach involves treating two comparable measures as orthogonal spatial dimensions, and to colour the voxel based on the angle formed between them in a circular colour space, based on the full 'HSV' colour map. The colours begin with red, pass through yellow, green, cyan, blue, magenta and return to red. The colour map was aligned so that pure cyan represents positive values in both dimensions. The opacity of the voxels is determined by the length of the vector formed by the two variables, therefore the magnitude of the relationship between the two variables determines how brightly coloured and visible a given voxel will be (**Figure 5.2**). This method of selecting voxel opacity has the effect of highlighting any voxel that is engaged by the stimuli and hiding any that are not. Strongly activated or deactivated voxels in either or both measures will be opaque, weakly activated voxels are more transparent. This colour mapping method is useful when comparing

equivalent measures between experiments as each measure gives an equal contribution to the colouring and transparency of each voxel.



**Figure 5.2.** Vector colour map.

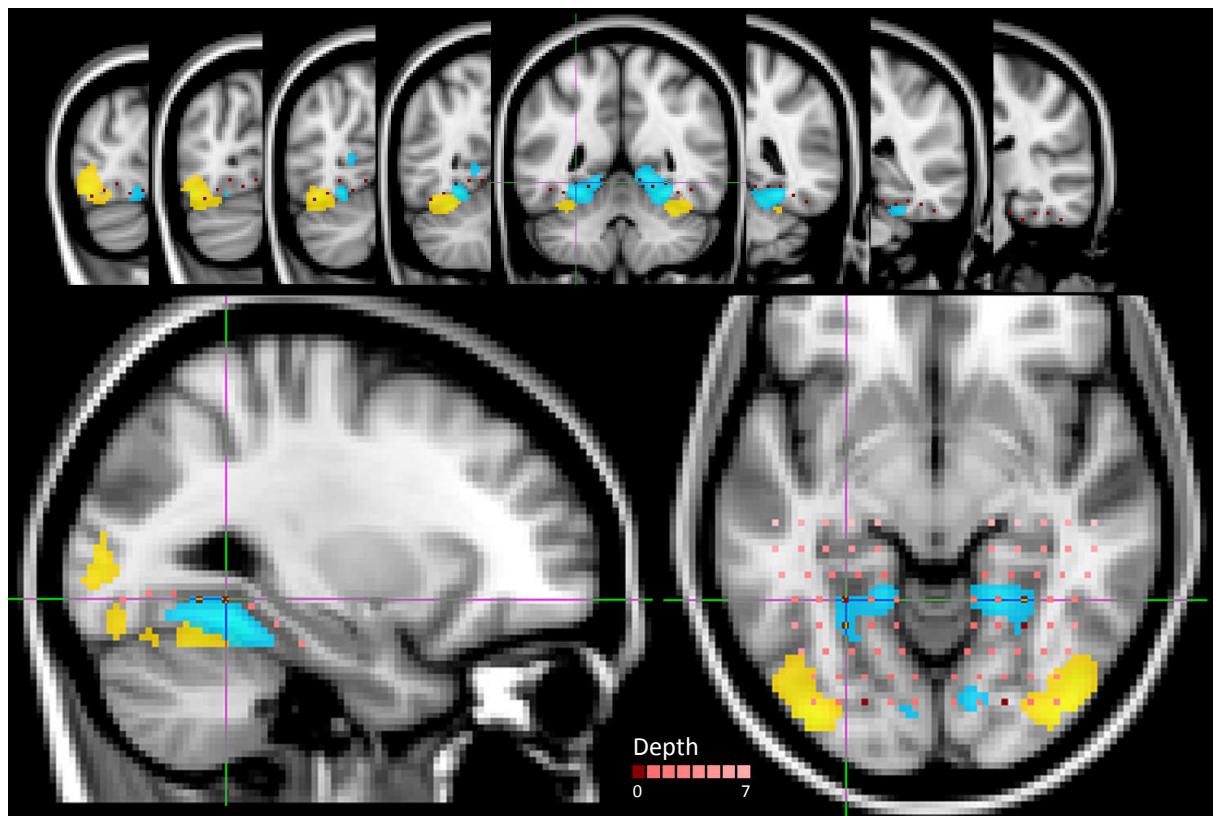
*Colour scale determined by the angle formed between two variables normalised between -1 and 1 (HSV colour map). Opacity determined by length of the vector formed by the two variables. This colour mapping technique can be applied to either selectivity or invariance measures.*

### 5.2.7 Spatial analysis

The colour mapping approaches described above are able to visualise the relationship between meaningful brain measures in terms of standardised effect sizes without the need for arbitrary statistical thresholds. They are able to reveal patterns of response that may be invisible to conventional whole brain mapping methods. It remains necessary however, to be able to test whether these patterns are non-random and reliable across individuals. To this end, a method of spatial sampling was devised whereby data could be extracted from each individual such that statistical comparisons could be conducted between not only different

measures from each of the experiments, but also different spatial locations in the brain. The spatial sampling was done in a manner independent of all functional measures, and based entirely on anatomical landmarks. An initial central line of points were defined in MNI space at evenly spaced intervals (8 voxels) along Y-dimension corresponding to the collateral sulcus in each hemisphere. The X and Z coordinates were manually adjusted so that each point remained within the collateral sulcus and fell within grey matter. The remainder of the grid was then defined from the points along the central line, by defining evenly spaced points (8 voxels) along the X dimension and adjusting the Z coordinates so that each point fell within grey matter (**Figure 5.3**). A total of 40 evenly spaced points were defined in each hemisphere.

The grid coordinates in MNI space were transformed to each individual native space using FLIRT (FMRIB's Linear Image Registration Tool) registration matrices calculated for each individual during each experiment's analysis. Data from the grid points were extracted from all of the functional brain images needed to compute measures of invariance and selectivity. All measures at the individual subject level were normalised into standardised units (Z-scores) using a grand mean and standard deviation across all grid points and subjects in a given experiment ( $[x^i - \text{mean}(x)] / \text{stdev}(x)$ ).



**Figure 5.3.** Cross sections of MNI brain spatial sampling grid.

All points defined relative to evenly spaced points in the Y dimension along the collateral sulcus. For display purposes sagittal and axial slices show all points, depth away from current slice denoted by red-pink colour map. Central slices centred on the MNI coordinates -28, -50, -8, and are displayed using neurological convention with the right hemisphere displayed on the right. Place (cyan, places>faces) and object (yellow, objects>scrambled) selective regions shown relative to grid points, defined using functional localiser data at the group level across all participants ( $n=69$ ).

### **5.3 Results 1: Object and scene view invariance**

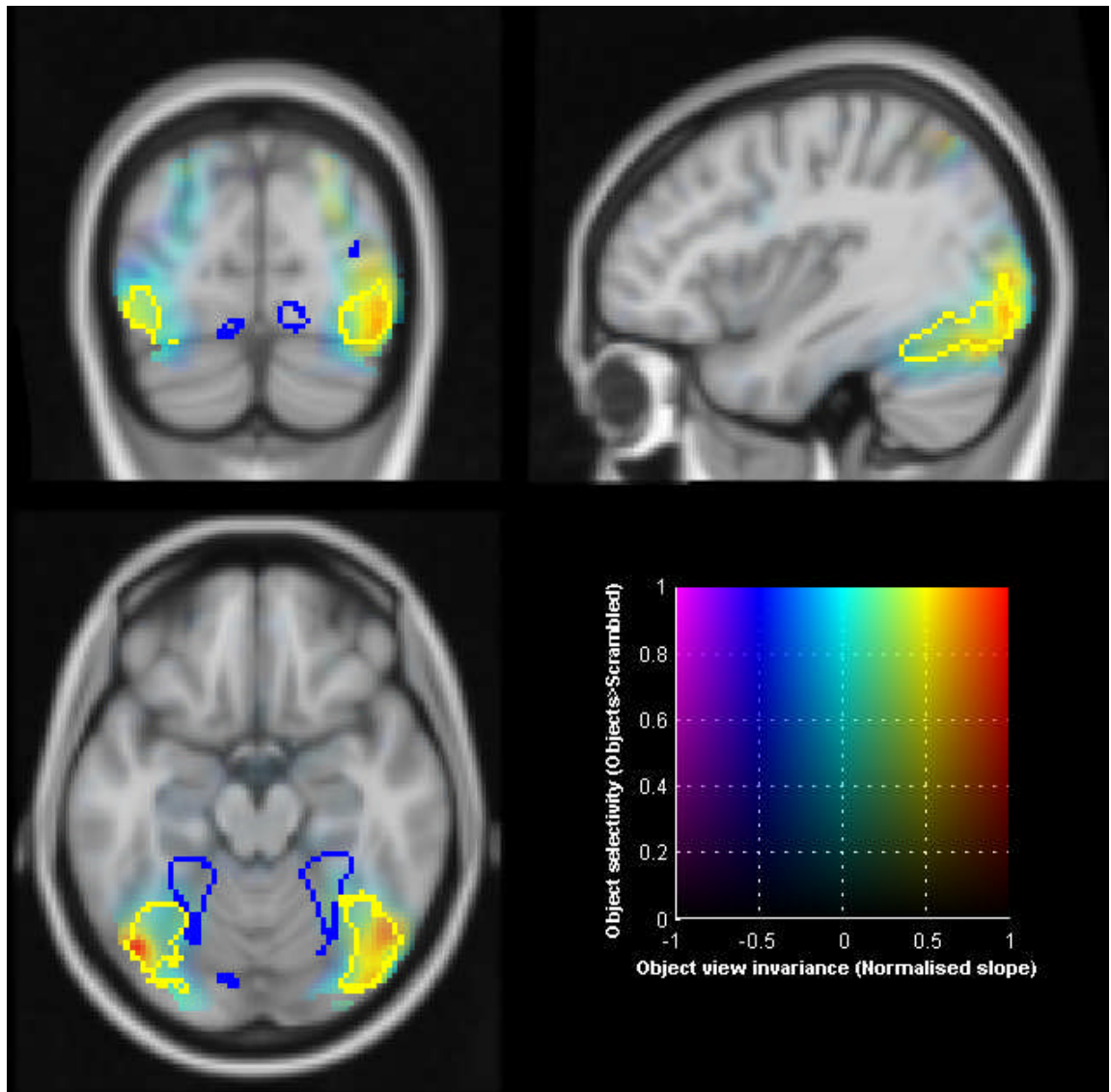
Chapter's 2 and 3 each examined how category selective cortical regions responded across parametrically manipulated view shift levels for landscapes and objects respectively. The view shift manipulations were the same in these two experiments, yet the responses across category selective cortex were very different. All category selective regions were found to be sensitive to view shift for scenes, whereas sensitivity to view shift for object was largely confined to object selective regions. Using the analysis and visualisation methods described here, it is possible to use comparable measures taken in these previous experiments and to create whole brain spatial maps of view invariance for both objects and landscapes and to investigate more generally how view invariance relates to ventral stream organisation. This section describes colour maps of invariance x selectivity (method section 5.2.5) for each of the above experiments, combines these two measures together for both invariance and selectivity measures using the vector colour mapping approach (method section 5.2.6), and these colour maps are described qualitatively. Spatial analyses are then conducted to show that the observed spatial patterns are meaningful and reliable, and differ depending upon stimulus type (method section 5.2.7).

**Figure 5.4** shows a whole brain group colour map of view invariance for objects, voxels coloured in red are highly view sensitive for objects, whereas voxels coloured cyan indicate view invariance. This colour mapping approach has the potential to reveal graded change in function as the intermediate colours between red and blue show varying degrees of view invariance/sensitivity. Voxel transparency reflects the object selectivity of that voxel with completely opaque voxels showing a high effect size in a contrast of objects>scrambled in the independent functional localiser scan. Voxels showing low responses in this contrast are transparent. Intermediate levels of object selectivity remain semi-transparent.

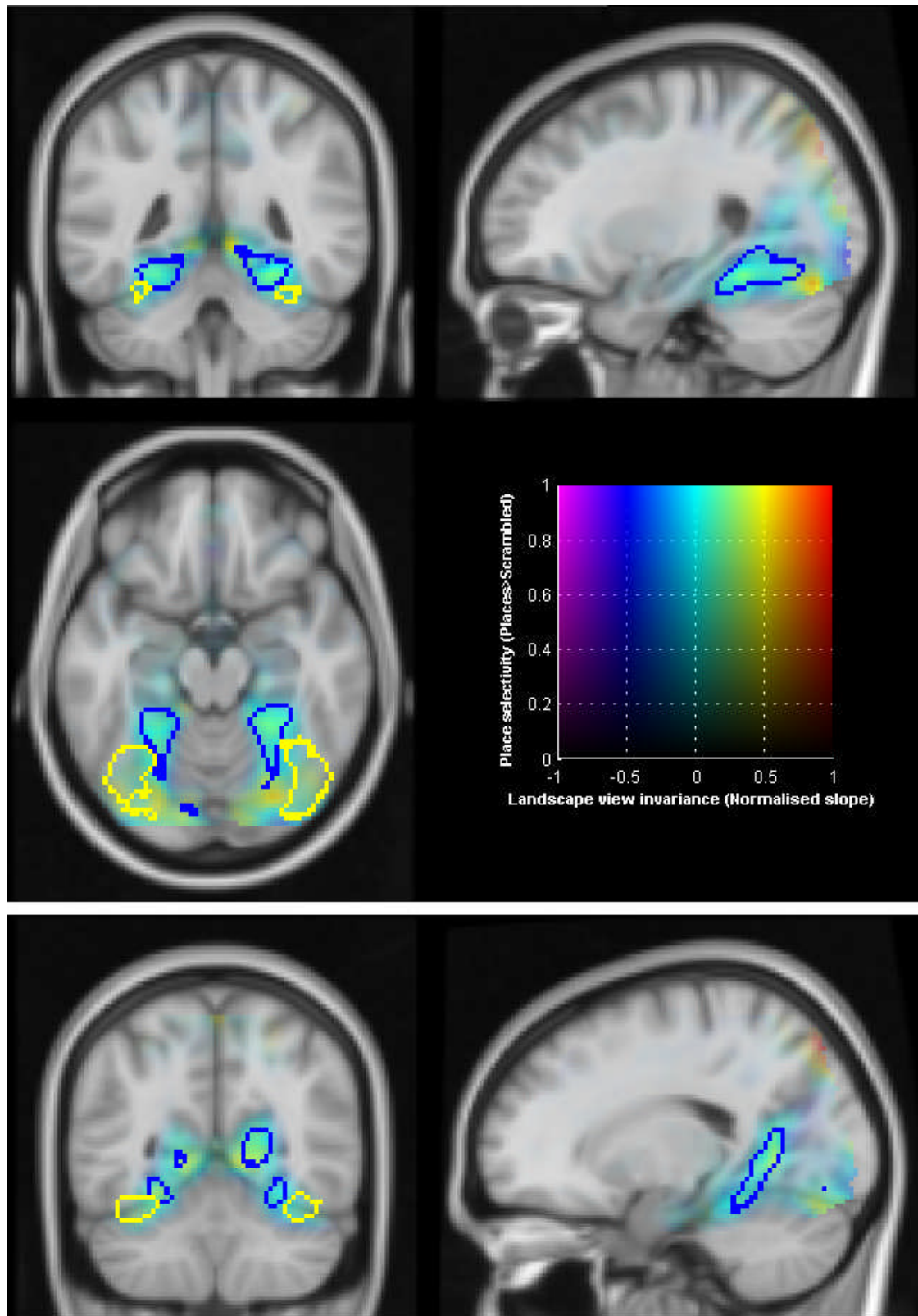
This approach makes it possible to see how invariance and selectivity relate to classical functionally defined regions without statistical thresholds. The borders

of place selective cortex (blue) and object selective cortex (yellow) are displayed on the figures. There are clear signs of functional distinctions within the borders of the object selective regions. The posterior part of the region shows a highly view dependent response, and towards the more anterior extent of the region there are signs that processing becomes increasingly view invariant. There are also signs of view invariant, yet object selective populations extending beyond the borders of object selective cortex. There is little sign of any object related processing within place selective cortex.

**Figure 5.5** shows the equivalent whole brain group colour map of view invariance for landscapes rather than objects, the same colour and opacity scheme is used to indicate levels of invariance and selectivity. The borders of place selective cortex (blue) and object selective cortex (yellow) are again displayed on the figure. Within the place selective region, the pattern of response appears to largely homogeneous, and predominantly view invariant, however, there is a clear indication of place selectivity and high levels of view sensitivity posterior to the place selective region, which correspond approximately to TOS/V3a (Nasr et al., 2011). There is an indication of an island of view sensitivity in the superior portion of the place selective region likely to correspond to the retrosplenial cortex (shown in **Figure 5.5**, additional, lower slices), a region implicated in spatial navigation and spatial memory (Vann et al., 2009). Whilst the retrosplenial cortex is place selective and is often localised using contrasts of places>faces, the functionally defined region is known not to correspond wholly to the anatomical region (Epstein, 2008; Vann et al., 2009; see chapter 2 for discussion). The present visualisation method is able to reveal aspects of scene processing in the retrosplenial region without being subject to the same limitations as conventional definitions of regions of interest.



**Figure 5.4.** Colour map of view invariance in object selective cortex. View invariance colour scale defined by normalised slope parameter of linear fit across four adaptation view shift conditions from the experiment in chapter 3. Voxels coloured red denote high levels of sensitivity, whereas voxels coloured cyan indicate an invariant response profile. Object selectivity determined by effect size measure in Objects>Scrambled comparison in functional localiser, and used to determine voxel opacity. Outlines of place (blue, places>faces) and object (yellow, objects>scrambled) selective regions defined at the group level using independent functional localiser scan. Slices centred on the MNI coordinates 38.29, -80.63, -17.76, and are displayed using neurological convention with the right hemisphere displayed on the right.



**Figure 5.5.** Colour map of view invariance in place selective cortex. View invariance colour scale defined by normalised slope parameter of linear fit across four adaptation view shift conditions from the experiment in chapter 2.



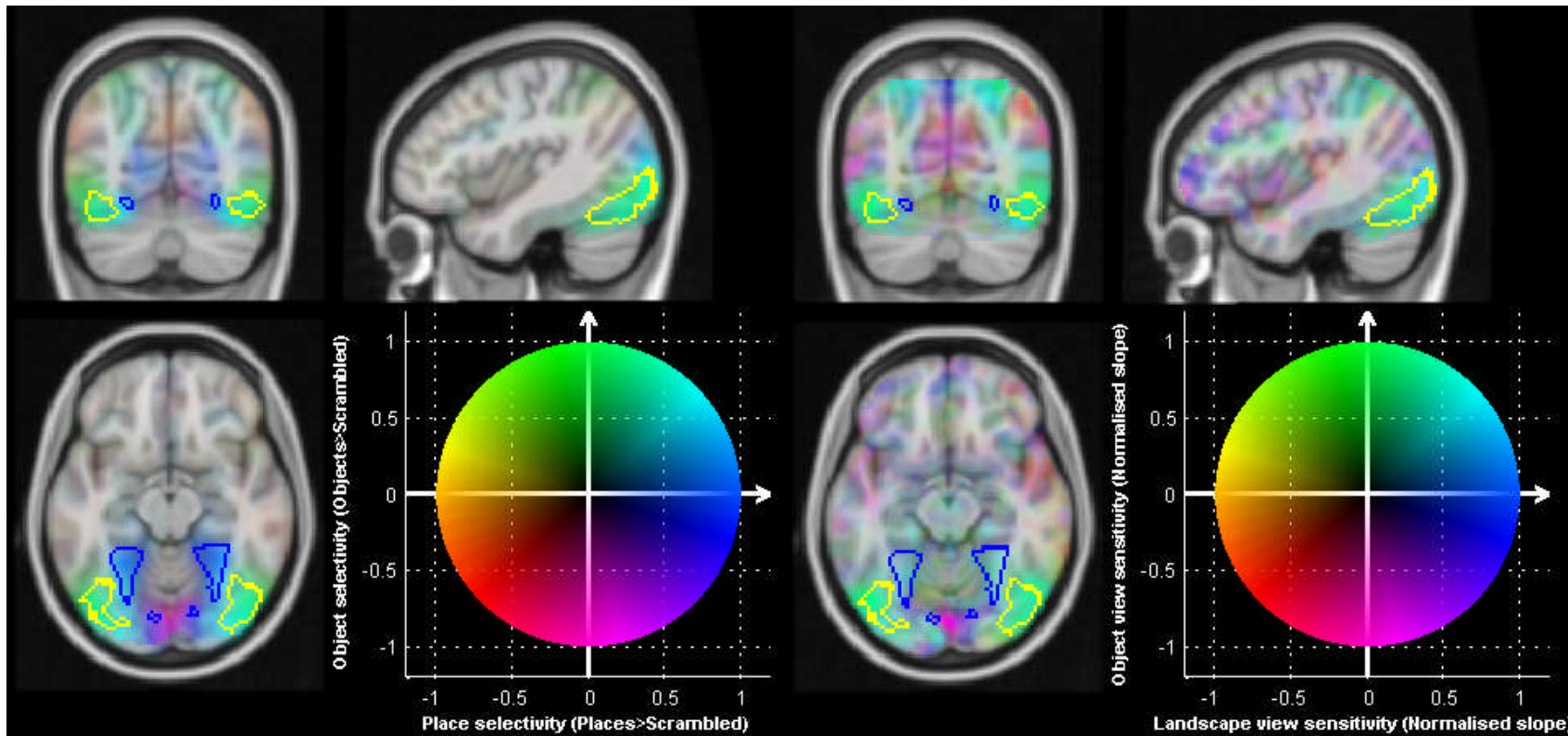
*Voxels coloured red denote high levels of sensitivity, whereas voxels coloured cyan indicate an invariant response profile. Place selectivity determined by effect size measure in Place>Scrambled comparison in functional localiser, and used to determine voxel opacity. Outlines of place (blue, places>faces) and object (yellow, objects>scrambled) selective regions defined at the group level using independent functional localiser scan. Upper slices centred on MNI coordinates 26.90, -45.84, -17.76, lower slices on 18.67, -57.22, 0.60. Images are displayed using neurological convention with the right hemisphere displayed on the right.*

### 5.3.1 *The relationship of object and scene processing*

In order to compare across the places and objects in measures of selectivity and invariance it is necessary to use the vector colour mapping approach (see **section 5.2.6**). The relationship of two comparable measures of effect size is visualised by colouring voxels based on the angle formed between them, and the voxels transparency determined by the length of the vector that they form. **Figure 5.6** shows the vector colour maps of the place and object selectivity measures and the place and object view sensitivity measures.

The leftmost cross section of **Figure 5.6** shows the relationship of object selectivity (objects>scrambled) to place selectivity (places>scrambled). Voxels coloured blue are primarily place selective, and voxels coloured green primarily object selective, with mixed selectivity represented with intermediate colour cyan. The opacity of the voxel is determined by the length of the vector formed between the two selectivity measures and shows all voxels activated or deactivated by either place or object stimuli. The map of selectivity largely corresponds to the functionally defined region borders for both stimulus types; there is some indication of multiple stimulus selectivity in regions posterior to the place selective border and within the posterior part of the object selective region. Additionally, some regions corresponding to low level visual cortex show a place selective response but a greater response for scrambled than intact objects (indicated by deep purple colouring). This is suggestive of regions that show a preferential response to low level properties of visual scenes and texture pattern stimuli (Haak, Renken, & Cornelissen, 2010).

The right most cross section of **Figure 5.6** shows the relationship of view sensitivity for landscapes to view sensitivity for objects. The method of selecting voxel opacity based on the length of the vector formed between the two view sensitivity measures has an important limitation for visualising invariance. The length of the vector in wholly invariant neural populations would be low and as such, invariant populations can be transparent. This method is only able to visualise the relationship between view sensitivity across the two studies. It is however, possible to describe unique patterns of response when considering the borders of category selective region and the invariance x selectivity colour maps from each individual study discussed above. Voxels within the place selective region are mostly transparent, yet are known to be highly place selective, this region appears to be largely invariant to view shift of places, and is not selective for objects. Posterior and lateral regions of object selective cortex appear to be sensitive to the view shift of both objects and landscapes. There is a region which appears sensitive only to object view shift towards the centre of the LOC, and the most anterior portion appears faint, suggestive of an invariant pattern of response when considering the equivalent location described by the invariance x selectivity colour map for objects.

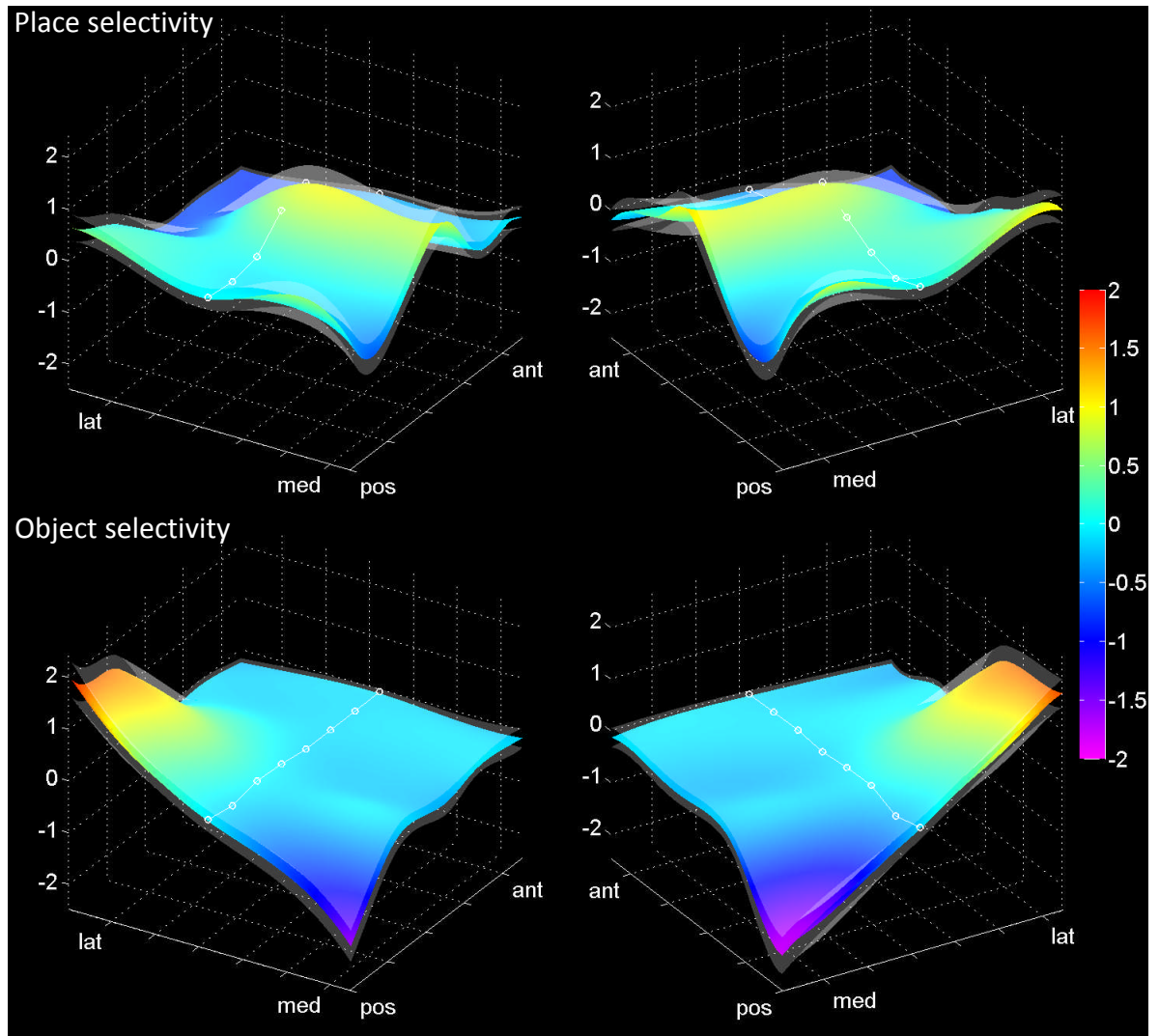


**Figure 5.6.** Vector colour maps of selectivity and view sensitivity.

Selectivity colour map (left) selects colour based on the angle formed between selectivity measures for objects (green) and places (blue), and opacity from the length of the vector formed by the two variables. View sensitivity colour map (right) uses the same colour and opacity procedure but using the view sensitivity measure (normalised slope) from the place and object adaptation studies (chapters 2, & 3). Outlines of place (blue, places>faces) and object (yellow, objects>scrambled) selective regions defined at the group level using independent functional localiser scan. Slices in both figures are centred on MNI coordinates -40.20, -67.40, -14.00 and are displayed using neurological convention with the right hemisphere displayed on the right

### 5.3.2 *Spatial analysis: selectivity*

In the preceding sections, the spatial patterns apparent in the whole brain maps were described qualitatively, the present section applies the statistical methods described in section **5.2.7** in order to compare the spatial patterns of scene and object selectivity, and to assess the reliability of the observed patterns. All localiser data across all prior experimental chapters were combined for a single within subjects analysis (n=69) in order to compare the spatial pattern of selectivity for objects and places. Data extracted from the grid points in each individual were averaged together and displayed as surface plots for the left and right hemisphere separately (**Figure 5.7**). Selectivity data from each of the experiments conducted were converted to standardised units (z-scores). A 2x5x8 (stimulus category, medial-lateral dimension, posterior-anterior dimension) repeated measures ANOVA was carried out. The analysis was carried out separately for each hemisphere to aid interpretability, although overall amplitude differences were found between the left and right hemisphere for selectivity measures. The results for all statistical comparisons of selectivity measures are shown in **Table 5.1**.



**Figure 5.7.** Place and object selectivity surface plots.

Selectivity measures from localiser contrasts places>scrambled (upper plots) and objects>scrambled (lower plots) extracted from anatomical MNI grid coordinates in z-score standardised units. Red colour denotes high levels of sensitivity, whereas cyan indicates an invariant response profile. Localiser measures across all three experiments used ( $n=69$ ). Data from left and right hemispheres are plotted separately on the left and right respectively. Surface between grid points generated using spline interpolation function. Error surfaces (white) above and below main surface display 95% confidence intervals. Collateral sulcus points marked in white along the central longitudinal axis of each surface.

**Table 5.1. Selectivity measures spatial grid ANOVA**

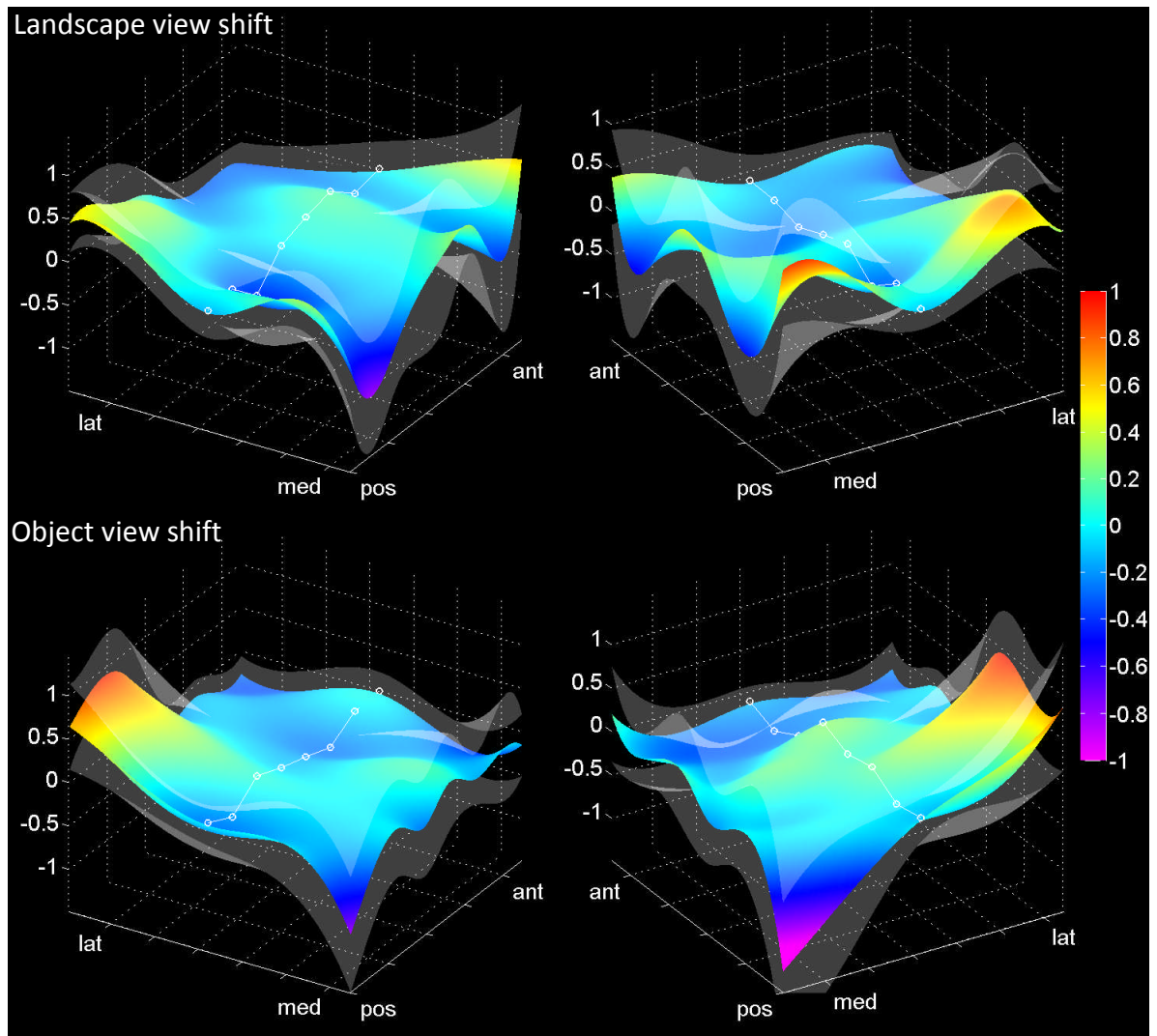
<b>Left Hemisphere</b>	
<i>Cat selectivity</i>	$F_{1,68}=13.183, P<0.05$
<i>Med-lat dimension</i>	$F_{4,272}=10.990, P<0.001$
<i>Post-ant dimension</i>	$F_{7,476}=37.551, P<0.001$
<i>Cat selectivity x med-lat</i>	$F_{4,272}=211.067, P<0.001$
<i>Cat selectivity x post-ant</i>	$F_{7,476}=40.818, P<0.001$
<i>Post-ant x med-lat</i>	$F_{28,1904}=39.531, P<0.001$
<i>Cat selectivity x Post-ant x med-lat</i>	$F_{28,1904}=24.416, P<0.001$
<b>Right Hemisphere</b>	
<i>Cat selectivity</i>	$F_{1,68}=12.830, P<0.05$
<i>Med-lat dimension</i>	$F_{4,272}=20.887, P<0.001$
<i>Post-ant dimension</i>	$F_{7,476}=23.574, P<0.001$
<i>Cat selectivity x med-lat</i>	$F_{4,272}=235.316, P<0.001$
<i>Cat selectivity x post-ant</i>	$F_{7,476}=42.599, P<0.001$
<i>Post-ant x med-lat</i>	$F_{28,1904}=53.254, P<0.001$
<i>Cat selectivity x Post-ant x med-lat</i>	$F_{28,1904}=17.466, P<0.001$

These results indicate that the spatial pattern of category selectivity throughout the ventral surface of the brain is reliably different between object and place stimuli. This is true along the posterior-anterior dimension, the medial-lateral dimension, and across the whole grid. Similar patterns are observed in both hemispheres. Lateral posterior peaks of object selectivity and medial anterior peaks of place selectivity are evident in the respective spatial grids

### 5.3.3 Spatial analysis: view invariance

Statistical methods described in section 5.2.7 are used to compare across object object- and landscape- view invariance experiments and assess the reliability of the patterns of sensitivity/invariance described in the whole brain maps. In order to compare the pattern of view invariance for objects and landscapes, a mixed design ANOVA was carried out, 2x5x8 repeated measures (hemisphere, medial-lateral dimension, posterior-anterior dimension), and a between subjects factor of experiment (landscape, object). Responses were largely consistent across hemispheres, main effects and interactions involving hemisphere were rare, and as such are only reported where significant. The surface data for each view invariance

measure are plotted in **Figure 5.8**, and the statistical comparisons are shown in **Table 5.2**.



**Figure 5.8.** Landscape and object view-invariance surface plots.

View invariance measures derived from the slope of linear fit across parametric landscape view shift conditions (upper plots) and parametric object view shift conditions (lower plots) extracted from anatomical MNI grid coordinates in z-score standardised units. Red colour denotes high levels of sensitivity, whereas cyan indicates an invariant response profile. Landscape experiment included 21 participants, object experiment included 22. Data from left and right hemispheres are plotted separately on the left and right respectively. Surface between grid points generated using spline interpolation function. Error surfaces (white) above and below main surface display 95% confidence intervals. Collateral sulcus points marked in white along the central longitudinal axis of each surface.

**Table 5.2.** *Landscape and object invariance spatial grid ANOVA*

<i>Experiment</i>	$F_{1,41} < 1, P=0.335$
<i>Med-lat dimension</i>	$F_{4,164} = 3.213, P < 0.005$
<i>Med-lat dimension x experiment</i>	$F_{4,164} = 3.127, P < 0.005$
<i>Post-ant dimension</i>	$F_{7,287} = 3.931, P < 0.001$
<i>Post-ant dimension x experiment</i>	$F_{7,287} = 2.956, P < 0.01$
<i>Hemisphere x post-ant x experiment</i>	$F_{7,287} = 2.478, P < 0.05$
<i>Post-ant x med-lat</i>	$F_{28,1148} = 4.688, P < 0.001$
<i>Post-ant x med-lat x experiment</i>	$F_{28,1148} = 3.054, P < 0.001$

These results indicate that the spatial pattern of view sensitivity throughout the ventral surface of the brain is reliably different for object and landscape view shift. This is true along the posterior-anterior dimension, the medial-lateral dimension, and across the whole grid. The lack of a main effect of experiment and the overall effects of the spatial dimensions suggest that observed differences are not due to overall differences between levels of activation. The pattern is generally flatter for landscape sensitivity with posterior peaks of high sensitivity in evidence. Object sensitivity is principally observable in lateral posterior portions of the VOT, with a generally flat response in anterior segments of the grid.

#### 5.3.4 *Discussion 1: view invariance in the ventral stream*

The results of chapters 2 and 3 revealed differences in the way category selective regions process view shift for their preferred and non-preferred stimuli. The present analysis was able to combine measures of both view-invariance and selectivity for landscape and object stimuli and go beyond this by revealing subtle spatial differences in object and landscape processing, both within the borders of category selective regions, and extending beyond them.

Considering selectivity maps in isolation, typical analyses of selectivity compare the responses of one stimulus type to that of another, and apply statistical thresholds to the resulting whole brain activation maps (Julian, Fedorenko, Webster, & Kanwisher, 2012; Saxe, Brett, & Kanwisher, 2006). Such approaches are



constrained to show discontinuous regions of activation with clear boundaries (Friston, Rotshtein, Geng, Sterzer, & Henson, 2006). The present method uses measures of effect size from conventional comparisons and is able to show symmetry and structure that would be hidden by conventional visualisation procedures. It also reveals graded change in selectivity within the borders of place and object selective regions, and voxels with intermediate selectivity for both stimulus types in areas immediately between them.

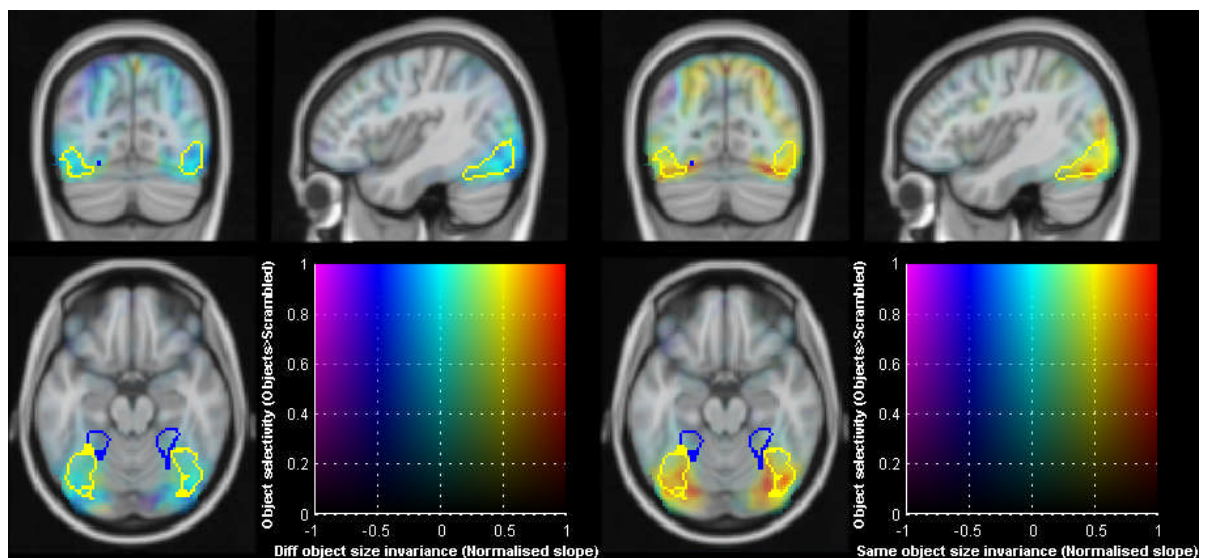
Considering measures of both invariance and selectivity, there is evidence of posterior/anterior gradients of high view sensitivity gradually becoming view invariant in more anterior voxels. This pattern is evident for both landscapes and objects, and is suggestive of distinct pathways for the two stimulus types. Object view-invariance is established in a graded fashion moving posterior-anterior and lateral-medial, whereas landscape view-invariance is established along a medial pathway and showing principally a posterior-anterior gradient. An organisation of gradually increasing invariance is widely believed to be present along the ventral stream (Grill-Spector & Malach, 2004; Ison & Quiroga, 2008), and is strongly suggested by single unit findings, where receptive field sizes gradually increase, and selectivity becomes increasingly complex (Kobatake & Tanaka, 1994; Rolls, 1994). The present patterns are broadly consistent with such hierarchical accounts of ventral stream processing, but suggest that the processing of different forms of visual change are systematically organised in the VOT, with distinct pathways associated with different stimulus types. The present method is able to use parametric adaptation to visualise such large scale patterns and show that such spatial patterns of view invariance and selectivity are statistically reliable and differ between stimulus category types.

#### **5.4 Results 2: Object Size Invariance**

In chapter 4, the parametric adaptation paradigm was modified to examine size invariance for objects. With manipulations of size shift, it was now possible to apply parametric manipulations to same identity and different identity objects. It is therefore possible to compute invariance measures independently of object identity. It was found that the response across size shift conditions within object selective cortex was modulated both by object identity and size shift. Object selective regions were found to be generally size dependent, however, it was also found that such patterns of size and identity modulation extend beyond the borders of functionally defined regions, and also vary over space within such regions. These findings are in accordance with the single unit literature which shows that both size invariant and size dependant neurons are present within monkey IT (Miyashita & Chang, 1988; Schwartz et al., 1983). Furthermore, single unit studies suggest that more anterior regions within IT, such as area TE, show greater size and position invariance (Ito et al., 1995; Lueschow et al., 1994). The present analysis is able to visualise size invariance at the population level, colouring voxels depending on the degree of size sensitivity, and show graded change in the response throughout the object processing stream.

**Figure 5.9** shows the whole brain colour maps of size invariance for objects (method section **5.2.5**). Voxels coloured towards the red end of the spectrum are highly sensitive to size changes, and voxels coloured cyan indicate a size invariant response. Voxel transparency reflects object selectivity, with completely opaque voxels showing a high effect size in a contrast of objects>scrambled in the independent functional localiser scan. Voxels showing low responses in this contrast are transparent. Intermediate levels of object selectivity are reflected in semi-transparent voxels. It is possible to see how size invariance and object selectivity relate to classical functionally defined regions, with the borders of place selective cortex (blue) and object selective cortex (yellow), displayed on the figure.

There are clear differences in size processing for same identity and different identity objects. When different objects are presented the response is largely invariant to size shift, and this pattern of response appears to extend beyond the borders of object selective cortex. When object identity is held constant there are islands of highly size sensitive responses, which again, extend beyond the border of the object selective regions in the posterior direction, however, there is also a trend for more size invariant response toward the anterior part of object selective cortex.



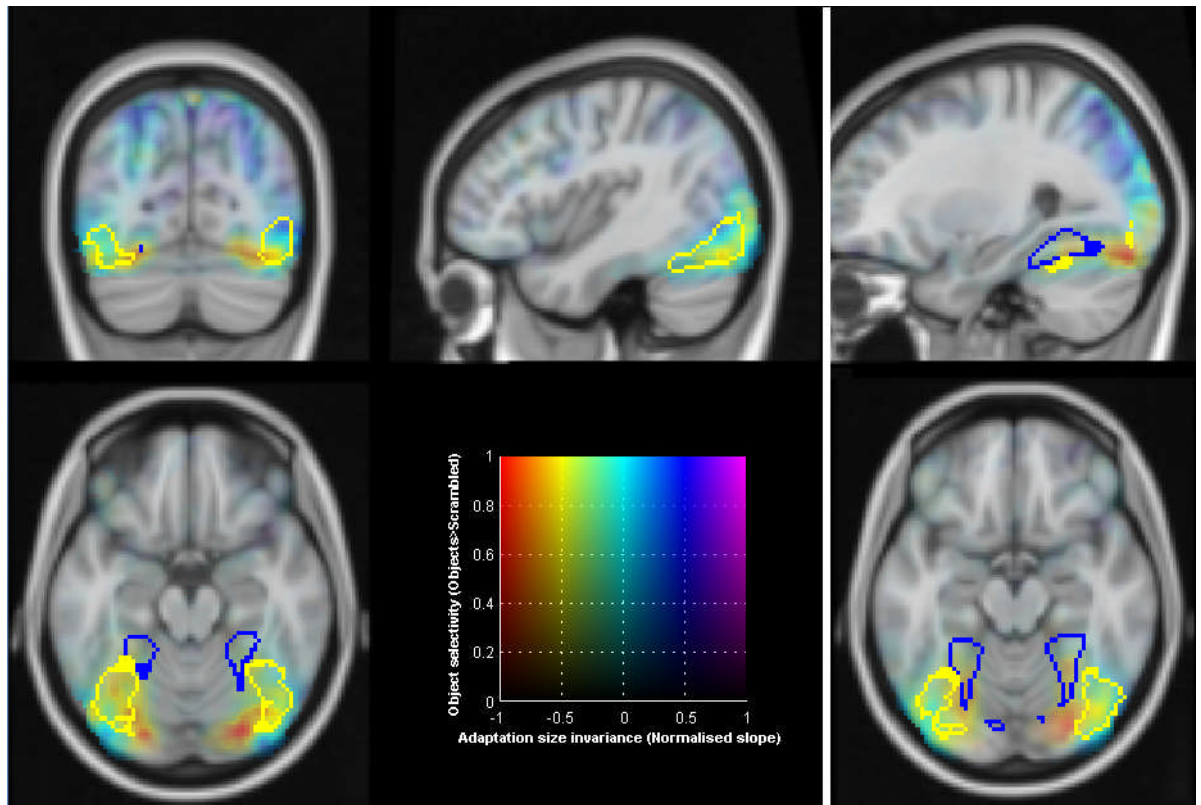
**Figure 5.9.** Colour maps of size invariance in object selective cortex.

Left slices show colour map for different identity objects, right slices show colour map for same identity objects. Size invariance colour scale defined by normalised slope parameter of linear fit across four adaptation size shift conditions from the experiment in chapter 4. Voxels coloured red denote high levels of sensitivity, whereas voxels coloured cyan indicate an invariant response profile. Object selectivity determined by effect size measure in Objects>Scrambled comparison in functional localiser, and used to determine voxel opacity. Outlines of place (blue, places>faces) and object (yellow, objects>scrambled) selective regions defined at the group level using independent functional localiser scan. Slices centred on MNI coordinates 41.45, -74.31, -19.02, and are displayed using neurological convention with the right hemisphere displayed on the right.

#### 5.4.1 Discrete baseline measure of size invariance

In the size invariance experiment (chapter 4), the addition of a second set of parametric size shift conditions for different identity objects allows a measure of

adaptation relative to a discrete baseline to be computed for each of the size shift conditions. It is therefore possible to calculate a slope across adaptation measures that can be considered an explicit measure of invariance driven by adaptation, where previously sensitivity measures necessitated remaining somewhat agnostic of the mechanism driving the response modulation. For each size shift condition a contrast of different>same was calculated, and using the effect size measures for this contrast, a linear function was fitted across the four adaptation conditions. The magnitude of adaptation is the inverse of brain response (adaptation measures expected to be high for small size shifts and low for large size shifts). The slope of a *negative* linear trend describes size- sensitivity/invariance. The colour scale has been inverted so that a negative slope is coloured red to reflect size shift sensitive responses and cyan to reflect size shift invariant responses. **Figure 5.10** shows the colour map generated using the adaptation slope measurement of invariance (method section **5.2.5**).



**Figure 5.10.** Colour map of discrete baseline size invariance in object selective cortex.

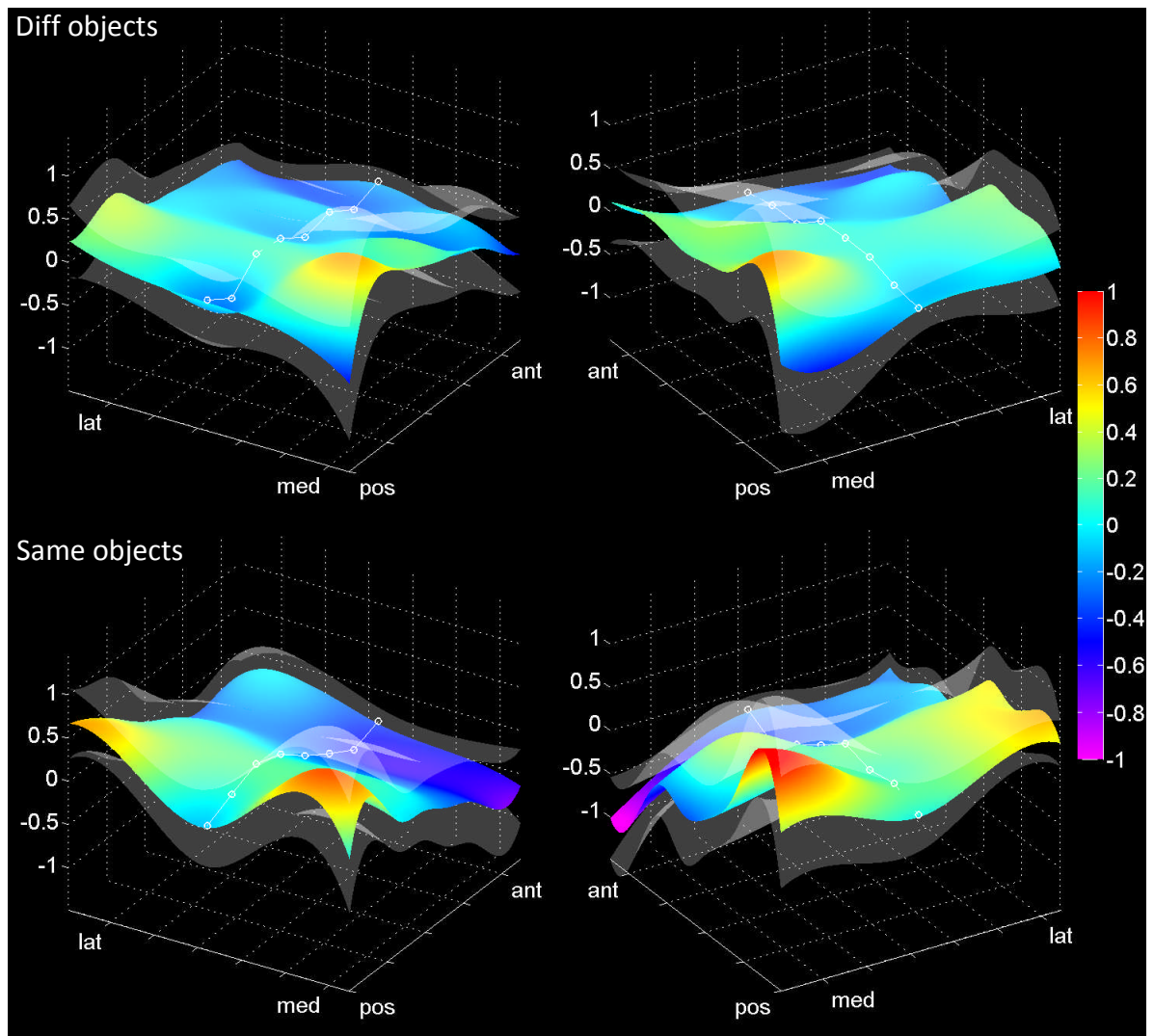
Size invariance colour scale defined by normalised slope parameter of a negative linear fit across four adaptation size shift contrasts (different>same) from the size invariance experiment (chapter 4). The colour scale has been flipped so that a negative slope is coloured red, reflecting size shift sensitive responses, and cyan reflecting size shift invariant responses. Object selectivity determined by effect size measure in Objects>Scrambled comparison in functional localiser, and used to determine voxel opacity. Outlines of place (blue, places>faces) and object (yellow, objects>scrambled) selective regions defined at the group level using independent functional localiser scan. Rightmost slices centred on MNI coordinates 41.45, -74.31, -19.02. Left slices centred on MNI coordinates 28.16, -59.12, -16.49. Images displayed using neurological convention with the right hemisphere displayed on the right.

The spatial profile of size shift adaptation shows gradients of size sensitivity to invariance throughout the extent of object selective cortex and beyond. Immediately posterior to the object selective region, there are populations of voxels showing size sensitive adaptation patterns. The more superior and anterior portions of object selective cortex show size invariant adaptation. A similar pattern of object

size sensitivity in posterior regions and size invariance in anterior regions can also be seen in place selective cortex (**Figure 5.10**, right).

#### *5.4.2 Spatial analysis: size sensitivity x identity*

Statistical methods described in section **5.2.7** are used to assess the reliability of the qualitative patterns of size invariance described above, and compare patterns between same- and different- identity objects. Sensitivity measures extracted from the grid points in each individual were averaged together and displayed as surface plots for the left and right hemisphere separately (**Figure 5.11**). These data are in standardised units and can be subject to within-subjects ANOVA in order to determine the statistical reliability of the patterns observed. In order to compare the spatial pattern of size sensitivity for different and same identity objects, a 2x2x5x8 repeated measures ANOVA was carried out (identity, hemisphere, medial-lateral dimension, posterior-anterior dimension). No main effects of hemisphere, or interactions involving hemisphere were found and as such statistics for hemisphere are not reported. The results for all statistical comparisons of size sensitivity measures are shown in **Table 5.3**.



**Figure 5.11.** Different and same object size-invariance surface plots.

Size invariance measures derived from the slope of linear fit across parametric different object size shift conditions (upper plots) and parametric same object size shift conditions (lower plots) extracted from anatomical MNI grid coordinates in z-score standardised units. Red colour denotes high levels of sensitivity, whereas cyan indicates an invariant response profile. Size invariance experiment included 26 participants. Data from left and right hemispheres are plotted separately on the left and right respectively. Surface between grid points generated using spline interpolation function. Error surfaces (white) above and below main surface display 95% confidence intervals. Collateral sulcus points marked in white along the central longitudinal axis of each surface.

**Table 5.3.** *Different and same object size invariance spatial grid ANOVA*

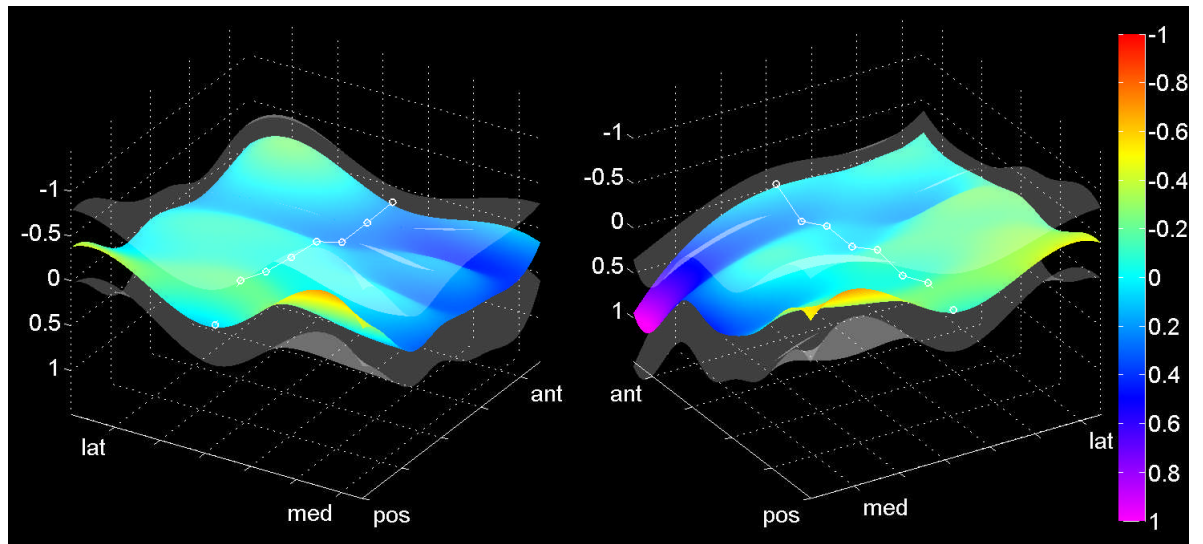
Identity	$F_{1,24} < 1, P=0.921$
Med-lat dimension	$F_{4,96} = 2.962, P=0.114$
Post-ant dimension	$F_{7,168} = 32.912, P<0.001$
Identity x med-lat	$F_{4,96} = 2.588, P<0.05$
Identity x post-ant	$F_{7,168} = 7.003, P<0.001$
Post-ant x med-lat	$F_{28,672} = 2.316, P<0.001$
Identity x post-ant x med-lat	$F_{28,672} = 1.617, P<0.05$

These results indicate that the spatial pattern of size sensitivity throughout the ventral surface of the brain is reliably different for same- and different- identity objects. This is true along the posterior-anterior dimension, the medial-lateral dimension, and across the whole grid. The lack of a main effect of object identity and the overall effects of the spatial dimensions suggest that observed differences are not due to overall differences between levels of activation. The pattern is much flatter for different identity objects, whereas same identity objects show a gradient of increasing size sensitivity towards more anterior parts of the VOT.

#### 5.4.3 *Spatial analysis: discrete baseline invariance measure*

Statistical methods described in section 5.2.7 are used to assess the reliability of the qualitative patterns of size invariance described in whole brain invariance maps. Discrete baseline adaptation data extracted from the grid points in each individual were averaged together and displayed as surface plots for the left and right hemisphere separately (**Figure 5.12**). In order to compare the spatial pattern of size invariance, a 2x5x8 repeated measures ANOVA was carried out (hemisphere, medial-lateral dimension, posterior-anterior dimension). No main effects of hemisphere, or interactions involving hemisphere were found and as such are not reported. The results for all statistical comparisons of same identity adaptation measures are shown in **Table 5.4**.





**Figure 5.12.** Object size-invariance surface plots (discrete baseline).

Discrete baseline adaptation measures reflecting size invariance, derived from the slope of linear fit across parametric adaptation conditions extracted from anatomical MNI grid coordinates in z-score standardised units. Z-axis and colour scale inverted as with adaptation measure, negative numbers indicate size sensitivity (red) and zero indicates size invariance (cyan). Experiment included 26 participants. Data from left and right hemispheres are plotted separately on the left and right respectively. Surface between grid points generated using spline interpolation function. Error surfaces (white) above and below main surface display 95% confidence intervals. Collateral sulcus points marked in white along the central longitudinal axis of each surface.

**Table 5.4.** Object adaptation size invariance spatial grid ANOVA

Med-lat dimension	$F_{4,100} = 3.207, P < 0.05$
Post-ant dimension	$F_{7,175} = 7.977, P < 0.001$
Post-ant x med-lat	$F_{28,700} = 1.692, P < 0.01$

These results indicate that the spatial pattern of size sensitivity as measured by discrete adaptation is reliable across individuals, and is apparent along the posterior-anterior dimension, the medial-lateral dimension, and throughout the whole grid. The general pattern is of highly size sensitive responses across the entire posterior part of the VOT, with a gradient towards invariant responses across the anterior part of the VOT.

5.4.4 *Discussion 2: size invariance in the ventral stream*

The present analysis shows how size sensitivity for different- and same- identity objects, and same-object, discrete baseline, sensitivity, change over space within object selective cortex and extending beyond it. This goes beyond what is possible both with conventional fMRI methods and single unit studies of size invariance. Voxels outside of, and posterior to, the LOC show highly size sensitive responses, this is also true of the posterior parts of LOC.

The discrete baseline conditions used to assess size change adaptation at each level of view shift, allows stronger inferences to be drawn from the resulting invariance maps. It is encouraging to note that the map of size sensitivity/invariance made with discrete baseline conditions, is not dissimilar to the map generated with same identity objects alone. This serves as a validation of the measure of sensitivity without discrete baseline conditions and suggests that maps generated from view shift manipulations can be considered a good approximation of maps driven by same identity adaptation.

Within the LOC, and towards more anterior ventral stream regions, there are islands of size invariant responses. The general pattern in lateral regions is similar to that of view invariance, with a posterior-anterior gradient of increasing size invariance. A gradient of size sensitivity to size invariance is apparent not only in object selective regions, but also along the medial pathway typically associated with scene processing. This suggests that size invariance may be established more generally throughout the VOT, and this is explored further in section 5.5, where maps of size- and view- invariance are compared directly.

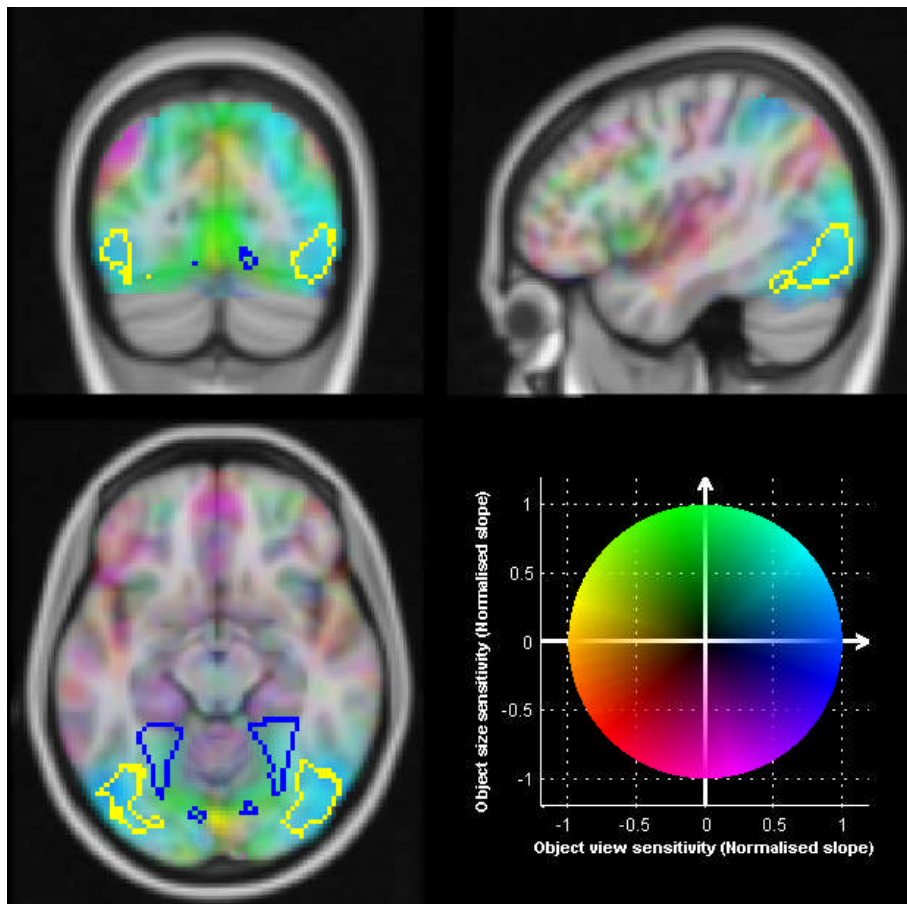
### **5.5 Results 3: Object size and view sensitivity**

Object size- and view- invariance measures were obtained in separate experiments using a closely related paradigm. The resulting measures of sensitivity-invariance across four same identity shift conditions are in each case normalised to fall between -1 and 1. In this sense the measures from the different experiment are comparable. This section investigates similarities and differences between the organisation size- and view-sensitive regions in VOT by combining data from chapters 3 and 4 in a single analysis. It is important to be clear that since size- and view-sensitivity relate to qualitatively different forms of visual change, overall differences in mean sensitivity may not be meaningful. However once such main effects of experiment are accounted for, interactions involving the spatial dimensions may be interpretable with caution. The aim is to investigate whether the spatial profiles of these two forms of invariance are dissociable.

The best way to visualise the relationship between these two measures is to use the vector colour mapping approach, where voxel opacity is determined by the magnitude of the vector formed by the two variables and voxel colour by the angle of the vector in circular colour space (see method section **5.2.6**). The vector length opacity selection has an important limitation for visualising invariance. The length of the vector in wholly invariant neural populations would be low, and as such, it is possible for invariant populations to be transparent. This analysis therefore focuses on the relationship of view sensitivity between the two studies. It is possible to reveal unique patterns of response when considering the borders of category selective regions and the individual invariance colour maps from respective sections discussed above.

**Figure 5.13** shows the vector colour map comparing object view and size sensitivity (see methods section **5.2.6**). The majority of object selective cortex appears to show size and view sensitivity in equal measure. The medial pathway, including place selective voxels and voxels outside of, and posterior to, category

selective regions, show size- but not view- sensitivity. The lateral part of the object selective region appears predominantly view-sensitive.

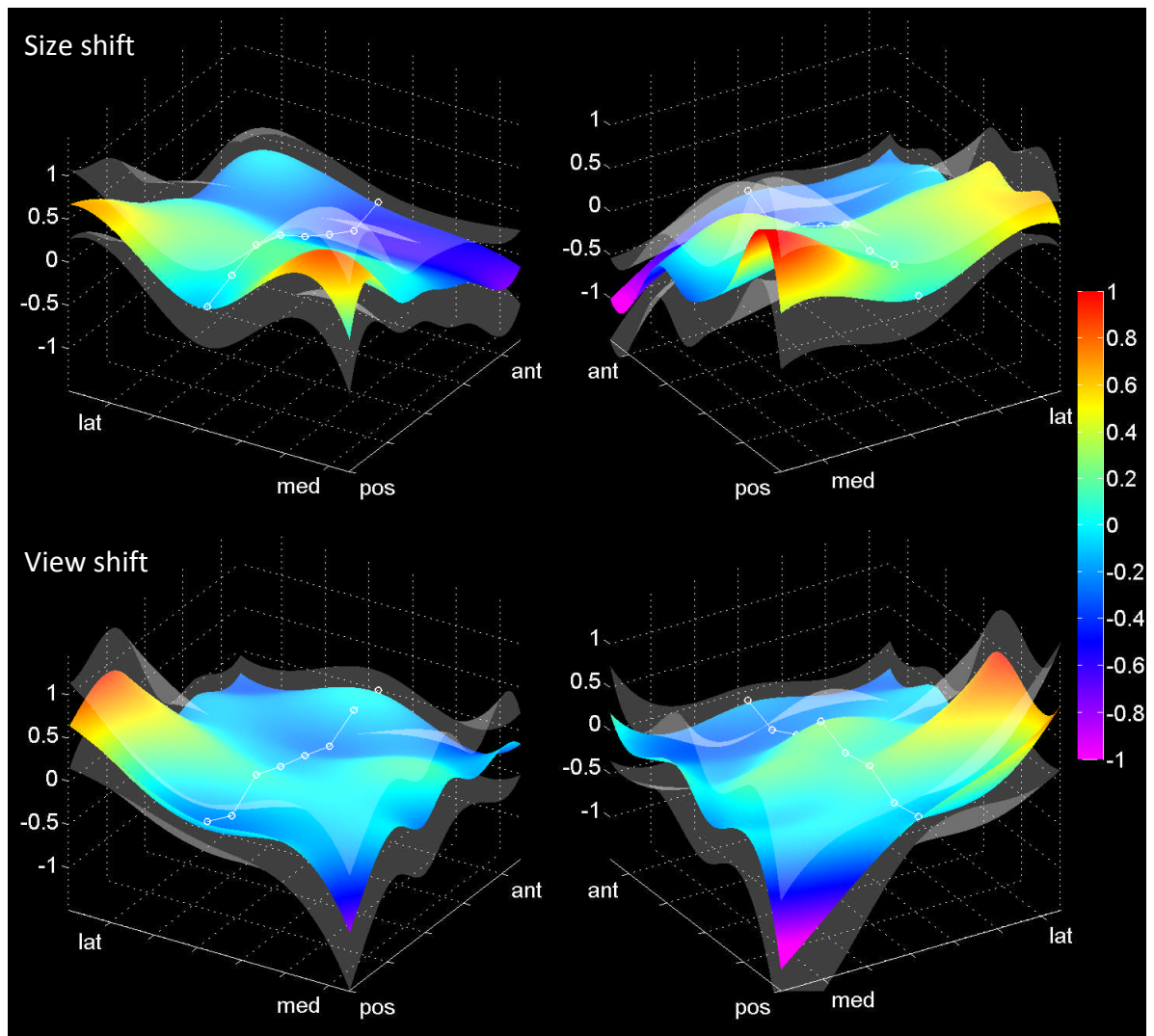


**Figure 5.13.** Vector colour map of object size and view sensitivity.

Voxel colour is selected based on the angle formed between view sensitivity (blue) and size sensitivity (green) measures for objects (normalised slope from adaptation scans, chapters 3 & 4). Opacity determined by the length of the vector formed by the two variables. Outlines of place (blue, places>faces) and object (yellow, objects>scrambled) selective regions defined at the group level using independent functional localiser scans. Slices are centred on MNI coordinates 43.35, -78.73, -13.33, and are displayed using neurological convention with the right hemisphere displayed on the right

*5.5.1 Spatial analysis*

Statistical methods described in section **5.2.7** are used to compare across object size- and view- invariance experiments and assess the reliability of the patterns described in whole brain maps. Sensitivity measures extracted from the grid points in each individual were averaged together and displayed as surface plots for the left and right hemisphere separately (**Figure 5.14**). These data from each of the experiments conducted are in standardised units (z-scores) and can be subject to within-, and between- subject ANOVA in order to determine the statistical reliability of the patterns observed. In order to compare the pattern of size- invariance to view- invariance for objects, a mixed design ANOVA was carried out, 2x5x8 repeated measures (hemisphere, medial-lateral dimension, posterior-anterior dimension), and a between subjects factor of experiment (object size, object view). The patterns were entirely consistent across hemispheres, there were no main effects or interactions involving hemisphere, and as such hemisphere statistics are not reported. The surface data for each invariance measure are plotted in **Figure 5.14**, and the statistical comparisons are shown in **Table 5.5**.



**Figure 5.14.** Object size- and view-invariance surface plots.

Size invariance measures derived from the slope of linear fit across parametric same object size shift conditions (upper plots), view invariance measures derived from the slope of linear fit across parametric same object view shift conditions (lower plots), extracted from anatomical MNI grid coordinates in z-score standardised units. Red colour denotes high levels of sensitivity, whereas cyan indicates an invariant response profile. Size invariance experiment included 26 participants, view invariance experiment included 22. Data from left and right hemispheres are plotted separately on the left and right respectively. Surface between grid points generated using spline interpolation function. Error surfaces (white) above and below main surface display 95% confidence intervals. Collateral sulcus points marked in white along the central longitudinal axis of each surface.

**Table 5.5.** *Object view- and size-invariance spatial grid ANOVA*

<i>Experiment</i>	$F_{1,46} < 1, P=0.277$
<i>Hemisphere</i>	$F_{1,46} = 5.042, P < 0.05$
<i>Med-lat dimension</i>	$F_{4,184} = 8.346, P < 0.001$
<i>Med-lat dimension x experiment</i>	$F_{4,184} = 1.495, P = 0.206$
<i>Post-ant dimension</i>	$F_{7,322} = 18.575, P < 0.001$
<i>Post-ant dimension x experiment</i>	$F_{7,322} = 5.070, P < 0.001$
<i>Post-ant x med-lat</i>	$F_{28,1288} = 4.488, P < 0.001$
<i>Post-ant x med-lat x experiment</i>	$F_{28,1288} = 4.931, P < 0.001$

These results indicate that the spatial pattern of sensitivity throughout the ventral surface of the brain is reliably different for object view- and size- invariance measures. This is true along the posterior-anterior dimension, and across the whole grid. The lack of a main effect of experiment and the overall effects of the spatial dimensions suggest that observed differences are not due to overall differences between levels of sensitivity. The posterior gradient of size sensitivity to size invariance is similar in medial and lateral segments of the grid. For view invariance this is evident only for the lateral extent of the VOT.

### 5.5.2 Discussion 3: separate forms of object invariance

The results of a comparison of spatial patterns of view- and size- sensitivity for objects suggest that these two forms of invariance are dissociable in the VOT. There is an apparent posterior medial-lateral gradient of size sensitivity to view sensitivity. Considering the medial and lateral pathways of the VOT, which are typically associated with scene processing or object processing respectively, the medial pathway is size sensitive, but not view sensitive, and the lateral pathway is sensitive to both size and view.

It is noteworthy that the observed gradients in view- vs. size- sensitivity parallels the postulated eccentricity biases described in the extrastriate visual system (Levy et al., 2001; Malach et al., 2002). Size change manipulations inevitably

differentially stimulate the visual field to varying extents. Therefore, the eccentricity bias affecting the organisation of VOT regions predicts that when objects fill the entire visual field they will tend to activate medial regions. The extent of visual field stimulation was tightly controlled in both size- and view- manipulations, such that throughout a single condition, the degree of visual field stimulation was on average, fixed. The size of the object stimuli used in chapter 3 was smaller than the largest sized objects used in chapter 4. It is possible that the view shift manipulation does not drive response to the same degree in peripheral parts of the map. Differences in the visual field properties of any view shift manipulation as compared to size manipulations may be inevitable, and may be inextricably linked with the pathways establishing different forms of invariance.

Establishing view invariance for objects appears to principally engage object selective regions, whereas size invariance for objects is established in medial and lateral regions. There is some suggestion that the parahippocampal gyrus is directly involved in processing navigationally relevant objects (i.e. landmarks) (Janzen & Turennout, 2004; Janzen, Wagensveld, & Turennout, 2007; Wegman & Janzen, 2011). The present findings suggest that the medial portion of the VOT, including parahippocampal gyrus, establishes size invariance for objects but does not establish view invariance. It is possible to speculate that during navigation, recognition of landmarks irrespective of their size/distance would be an important process, whereas recognition of landmarks from any viewpoint is less critical. The lateral pattern of response for both forms of object invariance is consistent with identity specific processing, building object representations that allow for fully invariant recognition (Milner & Goodale, 1995; Riesenhuber & Poggio, 2000).



## **5.6 Discussion**

Throughout this thesis, an experimental paradigm has been developed and refined that has been used to measure the extent to which various ventral stream regions are invariant, or sensitive to, high level stimulus changes. This approach has been applied to various stimulus types and different forms of invariance. The present chapter has outlined mapping techniques whereby measures of invariance can be defined at the level of an individual voxel and used to generate continuous whole brain maps that can reveal structure and symmetry beneath conventional statistical thresholds. There are various questions that can be asked about the way in which different kinds of invariance change over space throughout the ventral stream; various ways in which the different stimulus types, and different forms of invariance investigated throughout this thesis, can be compared with each other. The preceding sections within this chapter have addressed several of these questions.

Section 5.3 compared maps of view invariance for landscapes to view invariance for objects and showed that there are distinct pathways for the two stimulus types. Each of these pathways show separable posterior/anterior gradients of high view sensitivity gradually becoming view invariant towards more anterior voxels. Section 5.4 compared maps of size invariance between same and different identity objects, and combined these two measures to generate a map of invariant adaptation using discrete baselines for each level of size shift. It was apparent that size invariance may be a more general process than view invariance, with a gradient of increasing size invariance for objects apparent in both lateral and medial pathways within the ventral stream. Furthermore, the pattern for same object size sensitivity was qualitatively similar to the pattern for adaptation with discrete baseline conditions, suggesting that the measures previously taken for view shift are valid measures of same identity invariance. Section 5.5 further investigated the idea that size invariance may be a more general process by comparing spatial patterns of size sensitivity with patterns of view sensitivity. Apparent medial-lateral differences in how size and view invariance are established for objects was

revealed. The medial pathway was sensitive to size shift but not view shift for objects, perhaps suggestive of processing navigationally relevant object properties. The lateral pathway is sensitive to both size and view, suggestive of object identity specific processes.

### *5.6.1 Statistically assessing the reliability of spatial maps*

An aim of this chapter was to devise a statistical approach with which to assess the reliability of the spatial maps of selectivity and invariance. It is typical in cognitive neuroscience studies to attempt to draw inferences about the functional properties of brain regions from the results of fMRI experiments. However, the activation maps produced by conventional fMRI studies are the result of a comparison of the response in one experimental condition relative to another, it is rare to conduct comparisons between spatial points in the brain (Jernigan et al., 2003). The present mapping approach is able to show how the response properties of the ventral surface of the extrastriate visual system changes over space. In order to carry out spatial statistics, and assess the reliability of the observed patterns, grid points were defined in MNI space covering the majority of the ventral surface of the brain. Anatomical features such as the collateral sulcus and the gray/white matter boundary were used to define the grid, which was then used to sample from each individual's functional brain scans, each containing comparable measures transformed to standard units. This approach was able to assess the reliability of differences between the various different spatial maps. The measures of invariance described here are to some extent dependent upon the overall responsiveness of a particular brain region, which tends to reduce moving from posterior to anterior regions. The statistical approach adopted within this chapter, and used to compare between stimulus and invariance types, is able to show that maps of invariance and selectivity vary independently of large scale variation in responsivity in VOT.

There are several ways that the statistical approach for assessing the reliability of maps could be improved upon in future studies. The first is that the

grid of sample points in each individual is reliant upon a linear alignment to MNI stereotaxic space (FLIRT; FMRIB's Linear Image Registration Tool; Jenkinson et al., 2002). This could mean that in some cases the resulting sample points are less reliable than they ideally could be. This issue could be addressed by utilising recent advances in non-linear alignment algorithms (Klein et al., 2009). Alternatively the points could be anatomically defined in each subject's native brain space based on individual anatomy.

Another issue relates to the sampling of the ventral surface of the brain ignoring the convolutions of the cortex. Two points sampled close together in the volumetric coordinates may in reality be very far apart on the cortical surface. The ideal way to perform not just spatial sampling, but also visualisation of the ventral surface of the brain, would be to perform a surface reconstruction of the 2-dimensional cortical sheet (Dale, Fischl, & Sereno, 1999; Fischl, Sereno, & Dale, 1999; Wandell, Chial, & Backus, 2000). This approach is time consuming and computationally intensive, however, it could allow a more accurate depiction of the ventral visual processing hierarchy that can be independently defined in each individual before performing statistical comparisons.

### *5.6.2 Large scale organisation of the ventral stream*

An ongoing and central question that relates to ventral stream function is the extent to which category selective regions can be considered domain specific modules or continuous, overlapping maps (Op de Beeck et al., 2008). The present mapping approach is able to reveal symmetry and structure that is not typically visible using conventional fMRI approaches. Many previous studies may have been predisposed to the detection of functionally specific modules. Choices about particular stimulus manipulations, and the use of statistical thresholds, can influence the nature of the conclusions it is possible to draw. The present mapping imposes no restrictions on how graded or discontinuous the resulting maps will be. Qualitatively it is possible to see both smooth gradients and sharp changes in

function within and between category selective cortical regions. Future studies applying similar parametric adaptation methods and visualisation approaches to those used here could address questions about how continuous or discontinuous such changes in function are. It should be possible to do this quantitatively by considering measures of the rate of change in measures of invariance from voxel to voxel throughout the ventral surface. Such an approach would also be particularly powerful if applied in conjunction with cortical flattening techniques.

One of the central aims of this thesis has been to investigate large scale organising principles of the VOT. Conventional univariate approaches, are constrained to give categorical descriptions of the representations used by the brain (Op de Beeck et al., 2008). Multivariate approaches are able to show that a great deal more information is contained in signals measured throughout the VOT than had previously been thought (Haxby et al., 2001, 2011). Multivariate approaches however, are currently unable to tease apart and describe the dimensions of the representational space of the ventral pathway. The parametric adaptation paradigm combines measures of invariance and selectivity to high level properties of the visual stimulus in order to generate informative maps of the VOT and showing how invariance changes throughout the visual processing hierarchy.

The spatial patterns observed in the VOT are broadly consistent with hierarchical accounts of ventral stream processing (Grill-Spector & Malach, 2004; Ison & Quiroga, 2008), but suggest that the processing of different forms of visual change are systematically organised, with distinct pathways associated with different stimulus types and different forms of invariance. Distinct spatial patterns of parametric adaptation can be used to dissociate different forms of invariance. It is important to consider the different ways in which invariance is established when attempting to understand the large scale organisation of the VOT. Invariance varies by degree throughout the ventral surface of the brain depending on the nature of the stimulus, and the types of processing required.

## 6 Summary and general discussion

This thesis has investigated neural representations at the level of the VOT. A parametric adaptation paradigm was developed and used to investigate different stimulus types, and different forms of invariance associated with ventral stream processing. A key aim was to investigate how the process of establishing invariance influences, or is influenced by, the organisation of the VOT region.

Chapter 1 reviewed the current state of thinking regarding high level representations in the extrastriate visual system. It is now becoming clear that at this level in the visual processing hierarchy, the representational structure is best described as a multidimensional space within which visual stimuli can be located (Connolly et al., 2012; Dicarlo & Cox, 2007; Haxby et al., 2011; Op de Beeck et al., 2008). The response of a given region or voxel can be influenced by several stimulus dimensions at once and the nature of these dimensions is currently not well understood (Op de Beeck et al., 2008). The investigations conducted within this thesis were all grounded within this theoretical framework. Chapter 1 laid the foundation for an experimental paradigm with which individual dimensions of high level representation space can be investigated by parametrically manipulating different forms of visual change while measuring fMR adaptation. This paradigm was then applied and incrementally developed throughout later experimental chapters.

The first experimental chapter (chapter 2), investigated the representation of landscapes in place selective cortex. Previous studies had reported that the PPA was invariant to changes in view point (Ewbank et al., 2005), yet other studies, using greater changes in view, had reported that the representation of scenes in this region was view dependent (Epstein et al., 2003; Epstein, Higgins, et al., 2007). The parametric adaptation paradigm included multiple, linearly spaced, view shift conditions within a single experiment. Previous fMR-a studies measured adaptation as an all-or-none phenomenon, whereas the parametric adaptation approach

considers that adaptation, and by extension, invariance, may vary by degree. This experiment was able to show that the response in the PPA across view shift conditions is modulated by the degree of view shift. It was concluded that the representation of visual scenes in the PPA subserves the immediate perception of the visual world, establishing view invariance for degrees of change likely to be naturally encountered when moving through the environment. Invariant recognition across larger shifts in view, such as those that might be encountered in the process of topographical memory or navigation, is thought to be established by other scene selective regions including the RSC or the hippocampus (Vann et al., 2009).

Chapter 2 additionally investigated the RSC and hippocampus, the extent to which they show modulation by view shift, and also how the activity in these regions was related to topographical memory performance. It was found that both the RSC and bilateral regions within the hippocampus showed sensitivity to view shift. The response in the RSC was found to be strongly related to topographical memory performance, individuals who did poorly on the task showed a highly view sensitive response pattern and overall weak activation in this region, high performers showed the opposite pattern. The RSC therefore, was considered to be involved in establishing representations of landscape topography, invariant to large changes in viewpoint and other changeable aspects of landscapes, supporting an account of this region as transforming perceptual representations of the environment between egocentric and allocentric coordinate frames facilitating encoding and retrieval in memory (Byrne & Becker, 2007).

Chapter 2 established the viability of the parametric adaptation paradigm for the investigation of ventral stream, category selective, representations. A novel, shuffled control condition ruled out the role of predictability within blocks of view shifting stimuli (commonly used in fMR-a designs). A further finding of this experiment was that the LOC showed sensitivity to view shift for landscapes, and this formed the motivation for applying the parametric adaptation paradigm to investigate ventral stream representations of objects.

Chapter 3 set out to investigate view-sensitivity in the representation of objects (Ewbank et al., 2005; Grill-Spector et al., 2001; Kourtzi et al., 2003; Vuilleumier et al., 2002), and additionally, whether view sensitivity was a general property of category selective regions for all stimulus types and to investigate specifically whether object representations in the VOT are sensitive to the degree of view shift of object stimuli. It was found that the response in object selective LOC was modulated by the degree of view shift, furthermore, it was found that view sensitivity for object stimuli was restricted to object selective cortex, not found to be present in the PPA or throughout the VOT. Chapter 3 also set out to investigate the extent to which low level visual change within a block can account for adaptation at the level of category selective cortex (Andrews et al., 2010). A novel scrambled control condition was included whereby successive object images within a block were scrambled preserving the degree of change from image to image within in a view shift block. It was concluded that adaptation effects in object selective cortex are largely unaffected by low level visual change.

Considering the results across chapters 2 and 3, it can be concluded that there is a dissociation in how view shift is processed for objects and landscapes. View sensitivity for places is found for both and place- (medial), and object- (lateral) selective regions, suggestive of distinct pathways for establishing invariant representations for these two stimulus types.

Chapter 4 applied the parametric adaptation paradigm to investigations of size invariance for object stimuli (Ewbank et al., 2005; Grill-Spector & Malach, 2001; Grill-Spector et al., 1999; Sawamura et al., 2005). Manipulations of size afforded the opportunity to enhance the paradigm by including multiple discrete baseline conditions, whereby each level of size shift was carried out for both different and same object blocks. This was previously impossible for view manipulations due to the lack of a common reference frame across different objects/landscapes. In this chapter, size shift, and object identity were manipulated independently, and it was found that size sensitivity in object selective LOC was considerably greater when object identity was held constant. This size shift by identity interaction in LOC was

found not to be driven by low level visual/retinotopic factors as non-object selective regions (V1, PPA), were not modulated by object identity. The pattern of results in LOC were suggestive of an intermediate stage in computing size-invariant object representations (Abbott et al., 2008; Salinas & Abbott, 1997; Salinas & Thier, 2000).

Chapter 5 set out to go beyond conventional ROI and whole brain analysis methods and extend the parametric adaptation paradigm by quantifying invariance at the voxel level. The slope of the linear trend measure across shift conditions reflects the degree of sensitivity in a region or voxel, and the effect size of a comparison of intact with scrambled category stimuli obtained in independent localiser scans reflects the selectivity of a voxel. These two measures combined can be considered a metric of invariance. This chapter characterised invariance in this way and mapped changes over space throughout the VOT for different stimulus types and different forms of invariance. A novel visualisation method is described as well as statistical techniques for assessing the reliability of spatial patterns of invariance in the brain.

Chapter 5 brought together data from all of the preceding experimental chapters and considered comparable measures of different forms of invariance and selectivity from each of them. There are various questions that can be asked about the relationship between different forms of invariance, and how invariance is established for different stimulus types. Comparisons of measures of view invariance for objects and landscapes revealed that a graded hierarchical pattern is evident, but is suggestive of distinct pathways for the two stimulus types. Object view-invariance is established in a graded fashion moving diagonally, posterior-anterior and lateral-medial. Landscape view-invariance is established along a medial pathway showing a posterior-anterior gradient. It was possible to dissociate the spatial patterns for the two stimulus types, and this was true for both measures of selectivity and invariance. Further exploring the patterns of size-sensitivity and view-sensitivity for objects revealed that size invariance may be a more general process than view invariance. It was possible to dissociate the spatial patterns of size sensitivity and view sensitivity. View sensitivity was restricted to the lateral



pathway, however, both lateral and medial pathways showed gradients of size sensitivity. The spatial patterns observed in the VOT are broadly consistent with well known hierarchical accounts of ventral stream processing (Grill-Spector & Malach, 2004; Ison & Quiroga, 2008), but suggest that the processing of different forms of visual change are systematically organised, with distinct pathways associated with different stimulus types and different forms of invariance.

### **6.1 Hidden patterns of VOT organisation**

The mapping techniques developed in chapter 5 combined independent measures of invariance and selectivity in ways that provide a more complete description of function within VOT regions than conventional analysis methods. This can be illustrated by considering the results of chapter 2 which examined the representation of landscapes in place selective cortex, and the results of section 5.3, which visualised the parametric adaptation data and functional localiser measures of selectivity in whole brain colour maps of view invariance for landscapes. ROI results from chapter 2 showed that the place selective PPA was sensitive to changes in view shift, and the whole brain analysis of that chapter showed voxels posterior to the PPA that were view sensitive, and little sensitivity within its borders. Overall it was concluded that the representation within the PPA was view-invariant only to small changes in view.

A limitation of ROI analyses is that voxels showing sensitive and invariant patterns of response are aggregated together, and with the increased statistical power afforded by ROI methods it is possible to detect weak effects that might not appear in whole brain analyses. In some cases the mean response of a large region can fail to provide the best characterisation of function in that region. The whole brain linear trend analysis that was used to detect view-sensitive voxels also has an important limitation. Invariant voxels show a flat pattern of response across view shifts, and would fail to show effect when analysed in this way. As such, whole brain

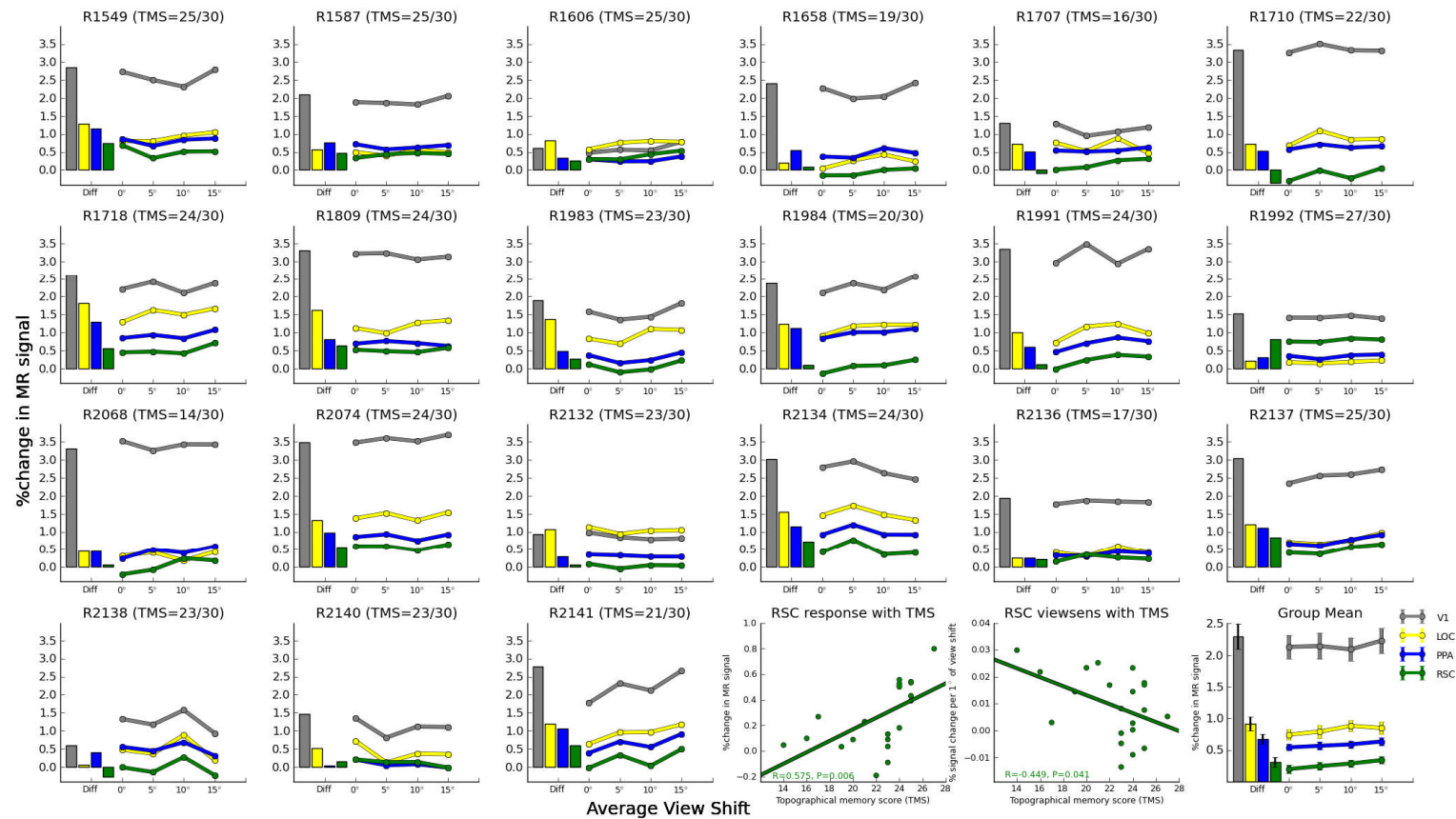
linear trend analyses are unable to differentiate view-invariant voxels from voxels that show no meaningful response to the stimuli.

Whole brain colour maps of landscape view invariance showed that the medial, place selective pathway gradually establishes view invariance moving posterior to anterior in the VOT. There are voxels that show sensitivity spanning the posterior borders of the PPA, but the majority of voxels within the region show a view-invariant response. At face value, these two analyses appear to give contradictory accounts, however, the whole brain colour mapping approach has the advantage that it simultaneously considers independent effect size measures describing selectivity and invariance, and is not subject to the same limitations of conventional ROI or whole brain measures. The mapping of parametric adaptation measures can reveal hidden symmetry and structure in the organisation of the VOT.

## **6.2 Conclusions**

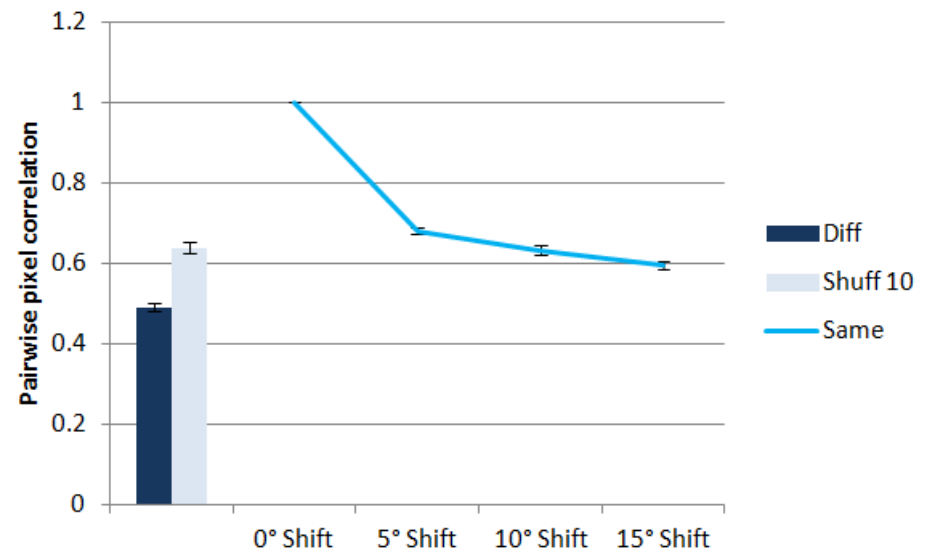
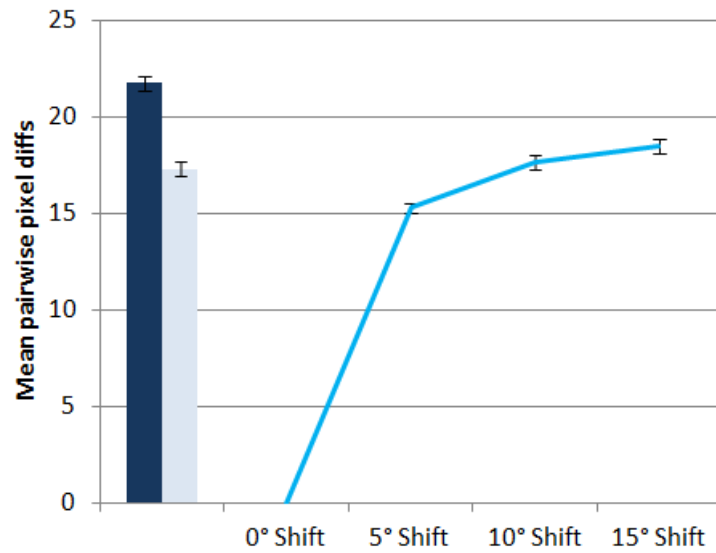
In order to fully understand the function of the ventral recognition pathway, it is necessary to characterise properties such as invariance and selectivity, and how they interact with the full range of visual properties involved in recognition. High level properties can be measured using parametric stimulus manipulations and this thesis shows how such measurements reveal structure and symmetry that can be hidden by conventional approaches. Invariance is a continuous variable that can be explicitly quantified rather than an all-or-none phenomenon. It varies by degree throughout the ventral surface of the brain depending on the nature of the stimulus and the types of processing required. Different forms of invariance can be dissociated from one another in terms of distinct spatial patterns of parametric adaptation that different forms of visual change generate.

## Appendix A



### Appendix A. 1. Landscape view invariance individual subject ROI subplots.

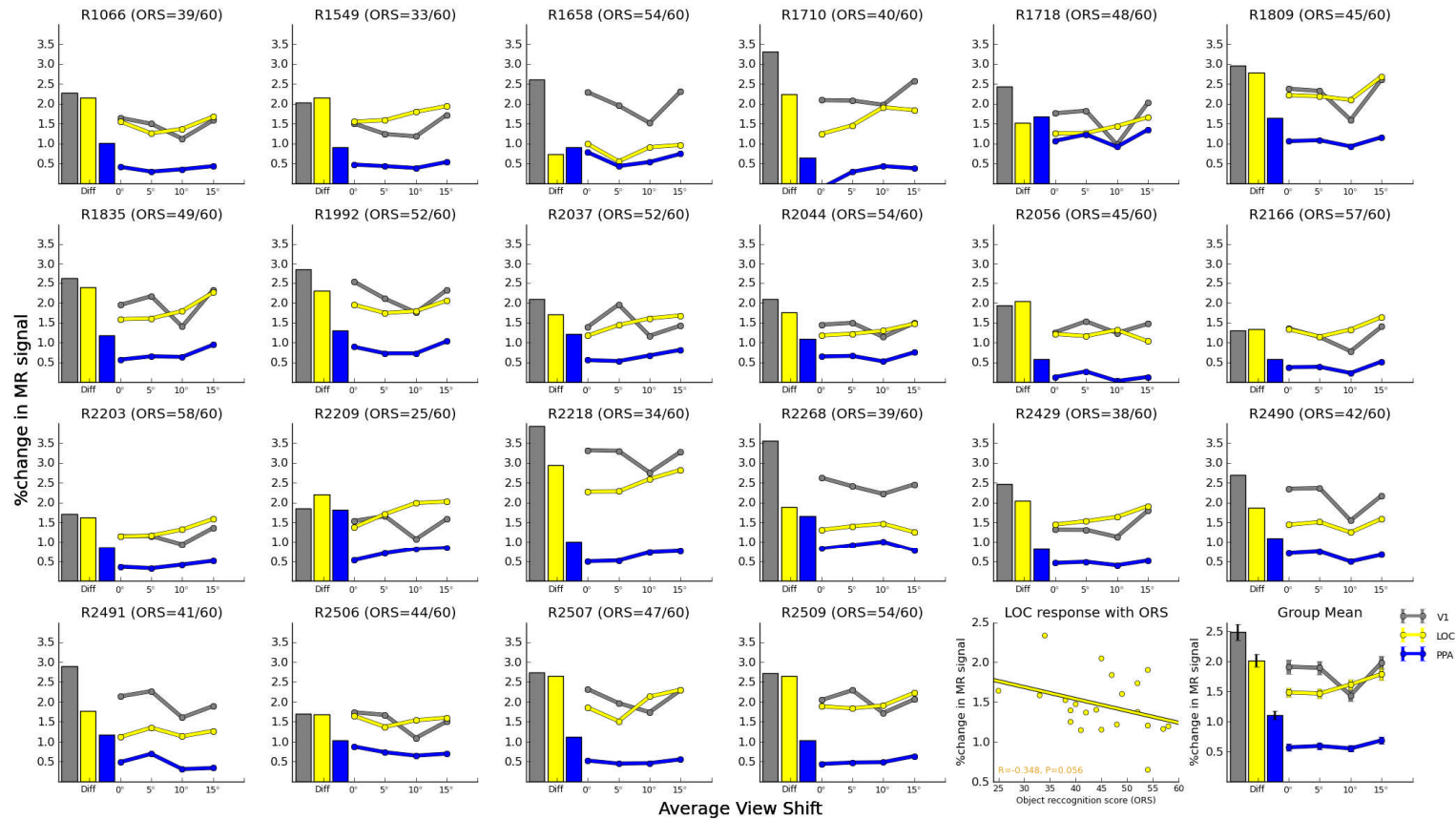
Each individual's ROI data contributing to the overall group Figure 2.3 along with each subject's topographical memory score (TMS) and group figures.



**Appendix A. 2.** Image level differences for Ch.2 landscape view invariance experiment.

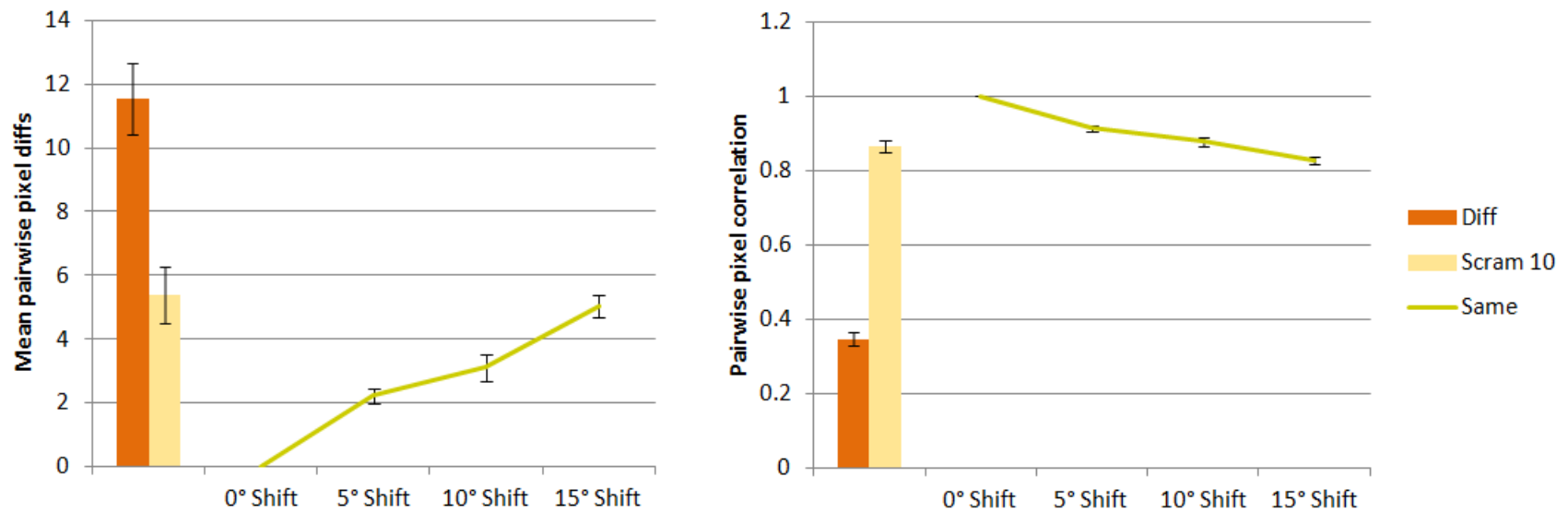
(Left) shows pairwise intensity differences across successive images within each block, averaged within conditions. (Right) Pairwise pixel level correlations across successive images in each block, averaged within conditions. Error bars represent  $\pm$  standard error of the mean.

## Appendix B



**Appendix B. 1.** Object view invariance individual subject subplots.

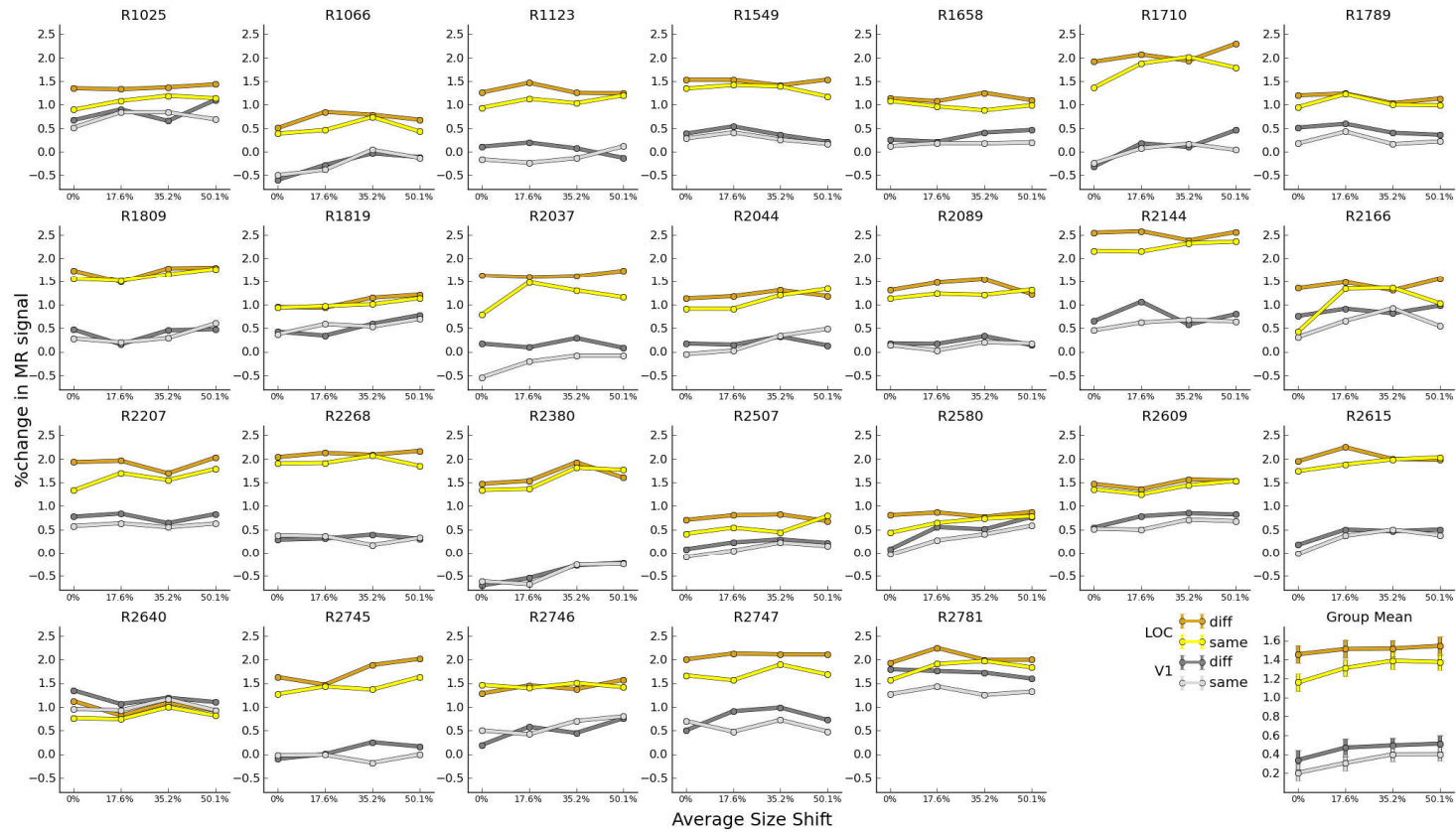
Each individual's ROI data contributing to the overall group Figure 3.4 along with each subject's object recognition memory (ORS) score and group figures..



**Appendix B. 2. Image level differences for Ch.3 object view invariance experiment.**

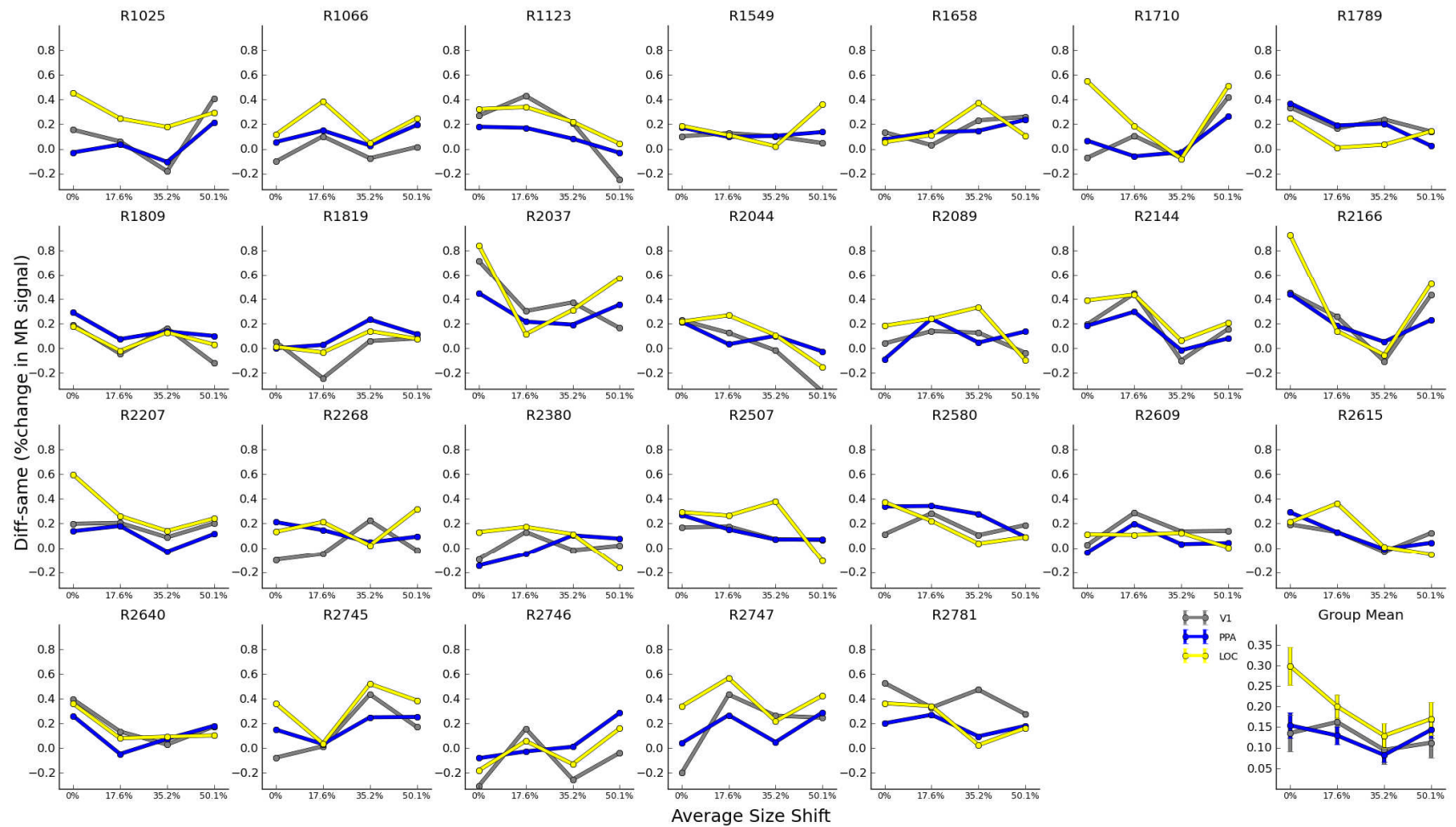
*(Left) shows pairwise intensity differences across successive images within each block, averaged within conditions. (Right) Pairwise pixel level correlations across successive images in each block, averaged within conditions. Error bars represent  $\pm$  standard error of the mean.*

## Appendix C



**Appendix C. 1.** Object size invariance individual subject ROI subplots.

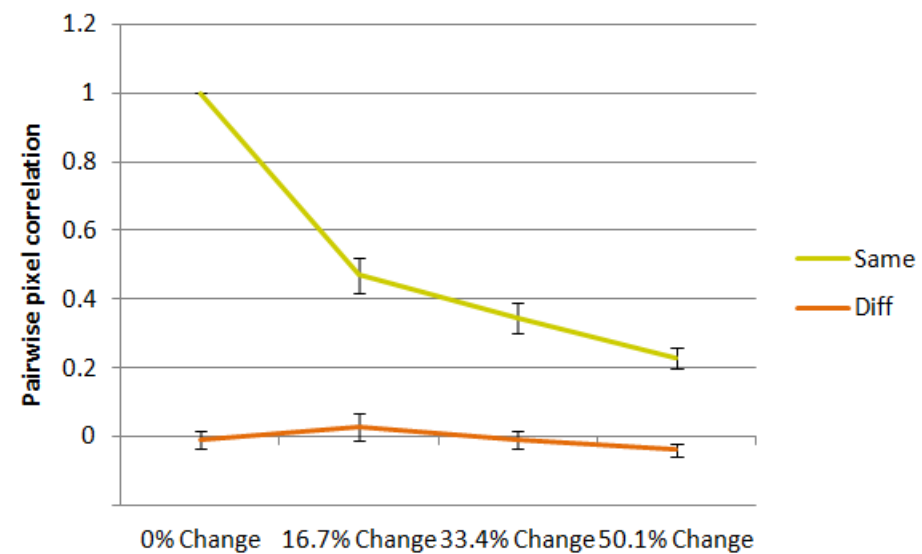
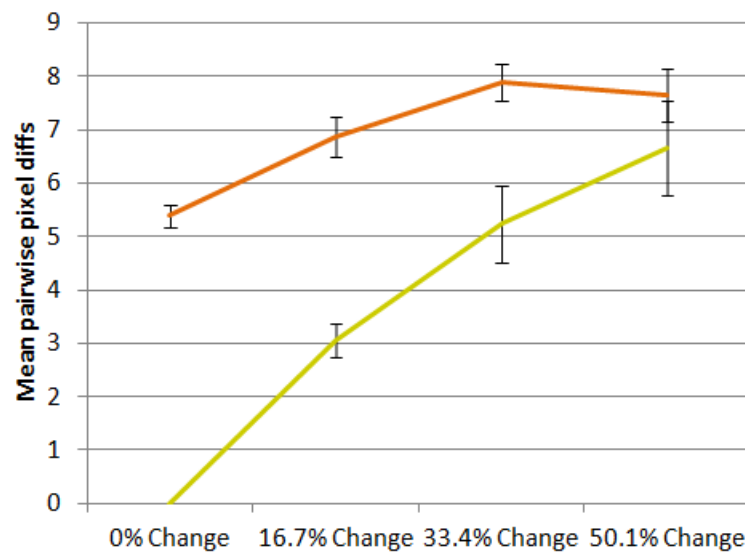
Each individual's ROI data contributing to the overall group Figure 4.6. BOLD response data from V1 and LOC regions of interest and group figure. PPA data not shown for clarity due to tendency to overlap with V1 data points (see Appendix C. 2).



**Appendix C. 2. Object size invariance individual subject ROI adaptation subplots.**

Each individual's ROI data contributing to the overall group Figure 4.7. Adaptation response data (diff-same) from all regions of interest and group figure.





**Appendix C. 3. Image level differences for Ch.4 object size invariance experiment.**

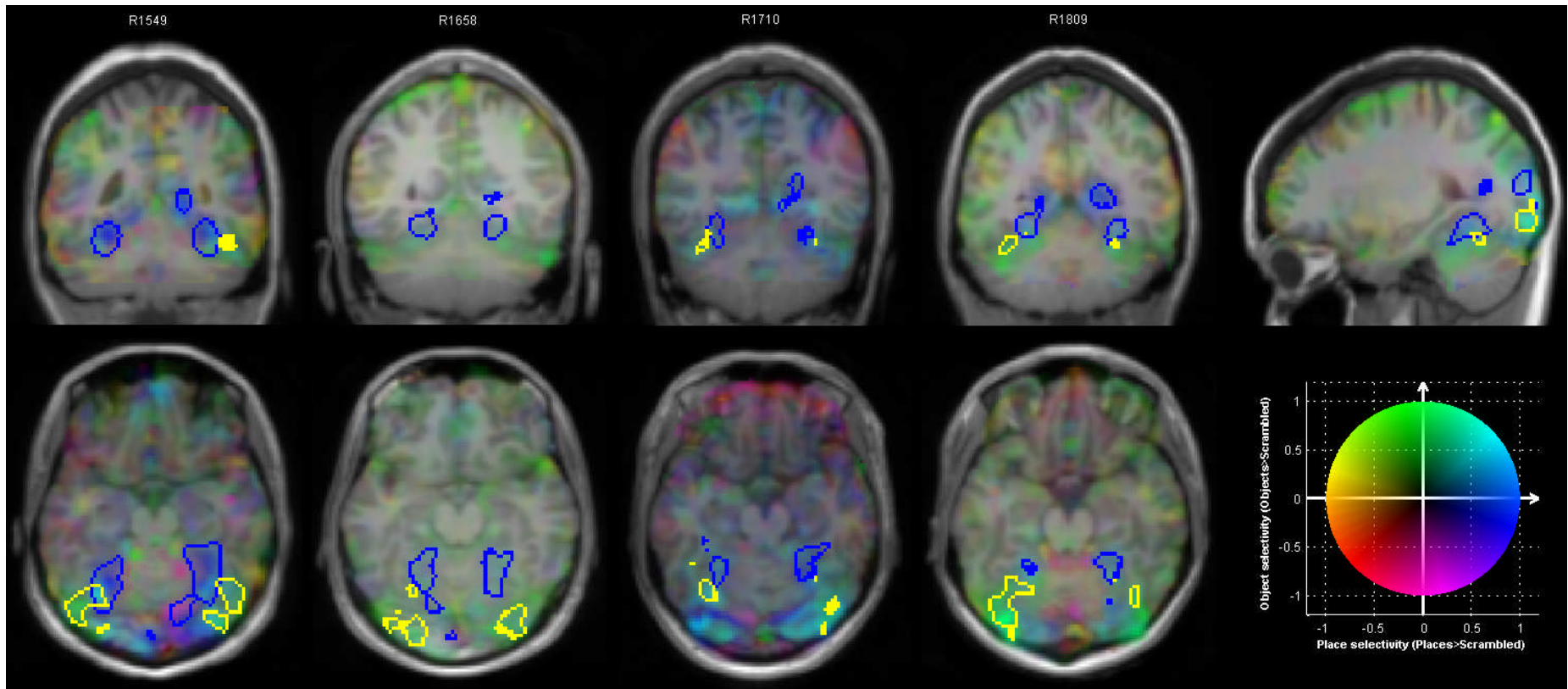
*(Left) shows pairwise intensity differences across successive images within each block, averaged within conditions. (Right) Pairwise pixel level correlations across successive images in each block, averaged within conditions. Error bars represent  $\pm$  standard error of the mean.*

## Appendix D

### *Within subject, across experiment analysis*

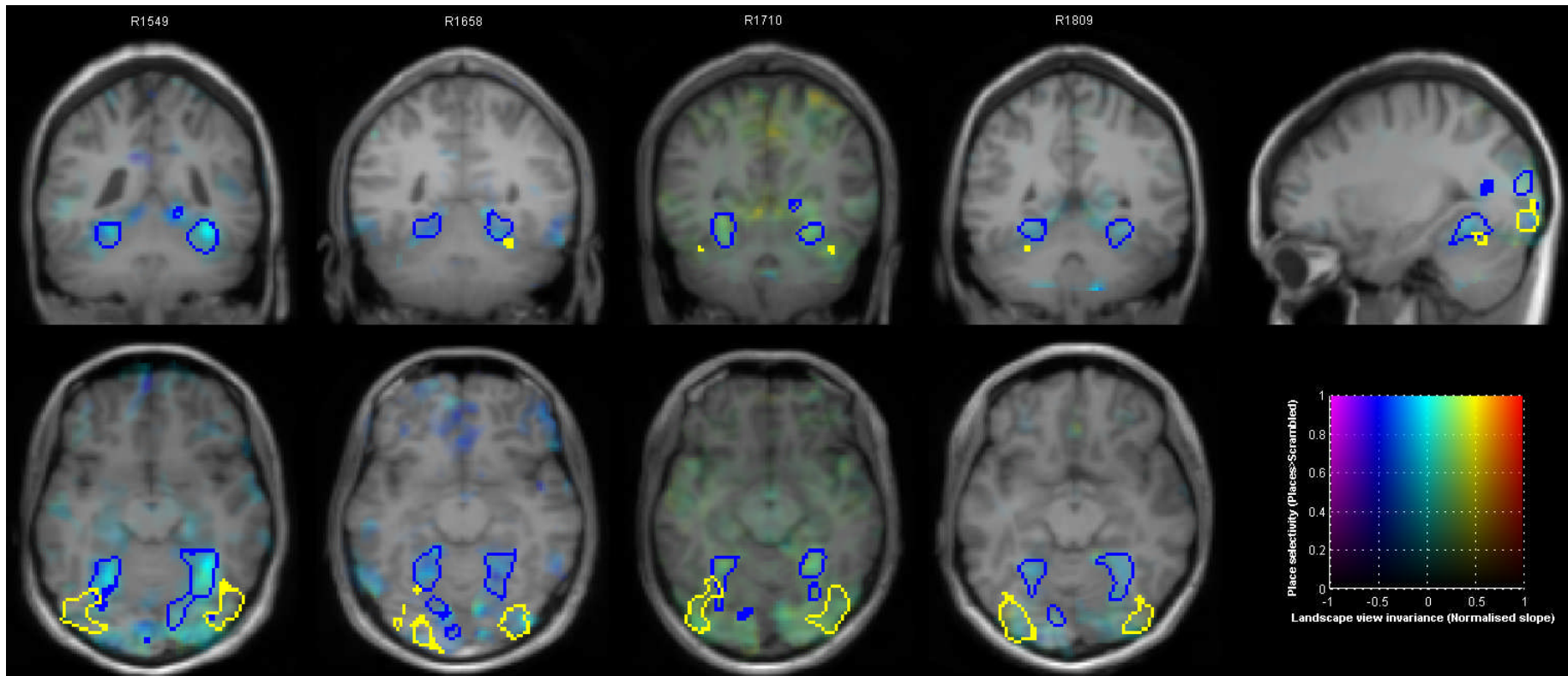
There were four individual subjects who took part in all three experiments in this thesis. This affords the opportunity to assess the reliability of spatial patterns of invariance and selectivity, and the measures created for them, in the brains of individual subjects. The present section reproduces the group colour maps described in chapter 5, but for individual subjects. It is then necessary to perform basic statistical comparisons to determine whether it is possible to discern different spatial patterns of invariance and sensitivity in small numbers of individuals. The main comparison that can be made across all experiments is between the patterns of responses across parametrically manipulated same identity stimuli.

For each subject, whole brain images containing measures of invariance and selectivity, and each participant's T1 structural scan, were transformed into MNI space using FLIRT (FMRIB's Linear Image Registration Tool) registration matrices calculated for each individual during each experiment's analysis. For the localiser data, a vector colour map was produced for each subject showing the relationship between object and place selectivity. **Figure D.1** shows vector colour map slices from each subject all centred on the same MNI coordinates. For each of the three adaptation experiments, the same identity invariance colour map for that experiment was recreated for each of the four subjects (see methods section **5.2.5**). **Figure D.2, D.3, & D.4**, show slices from each subject all centred on the same MNI coordinates for each of the three invariance experiments respectively. Although there is individual subject variability, similar qualitative patterns observed at the group level are apparent in individuals.

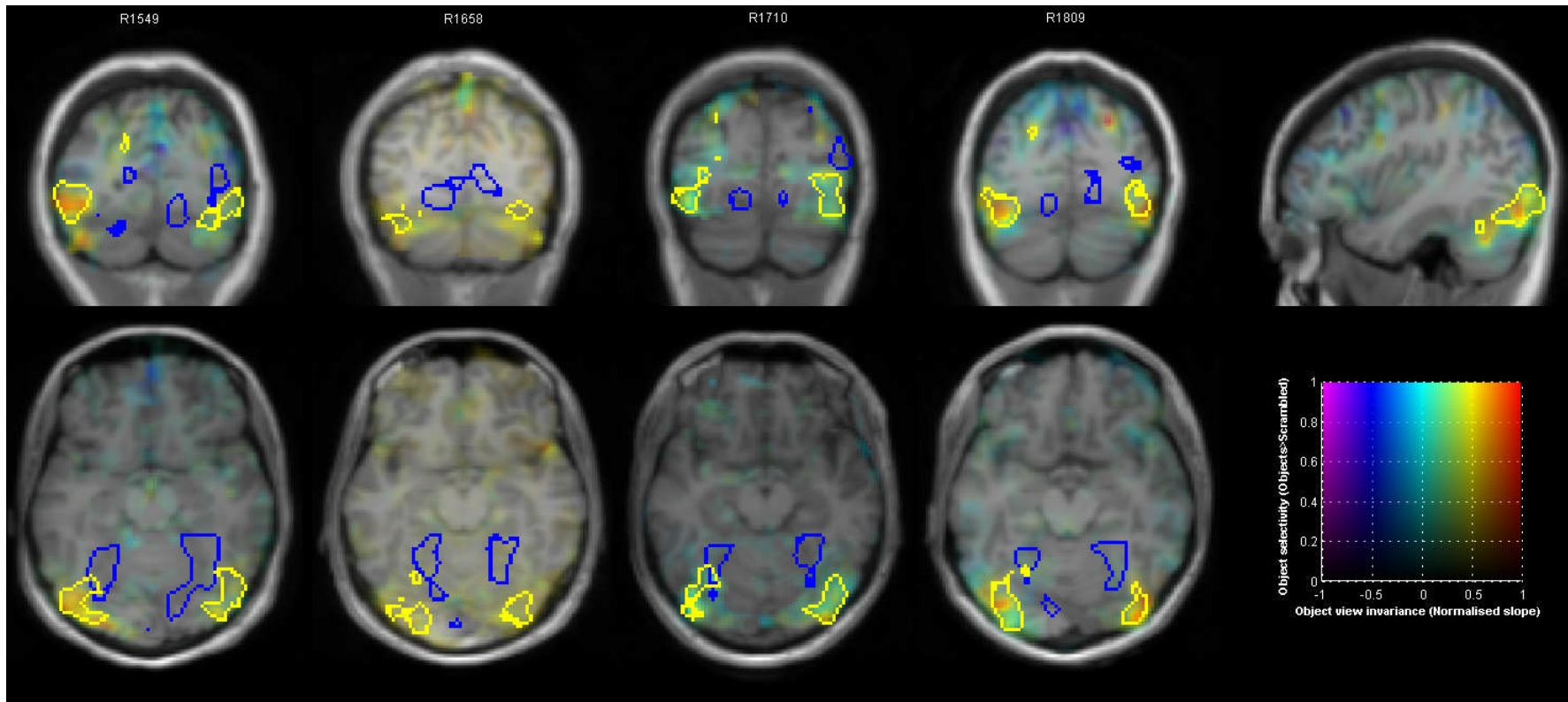


**Figure D.1.** Vector colour maps of category selectivity for four individual subjects transformed to MNI space (FLIRT).

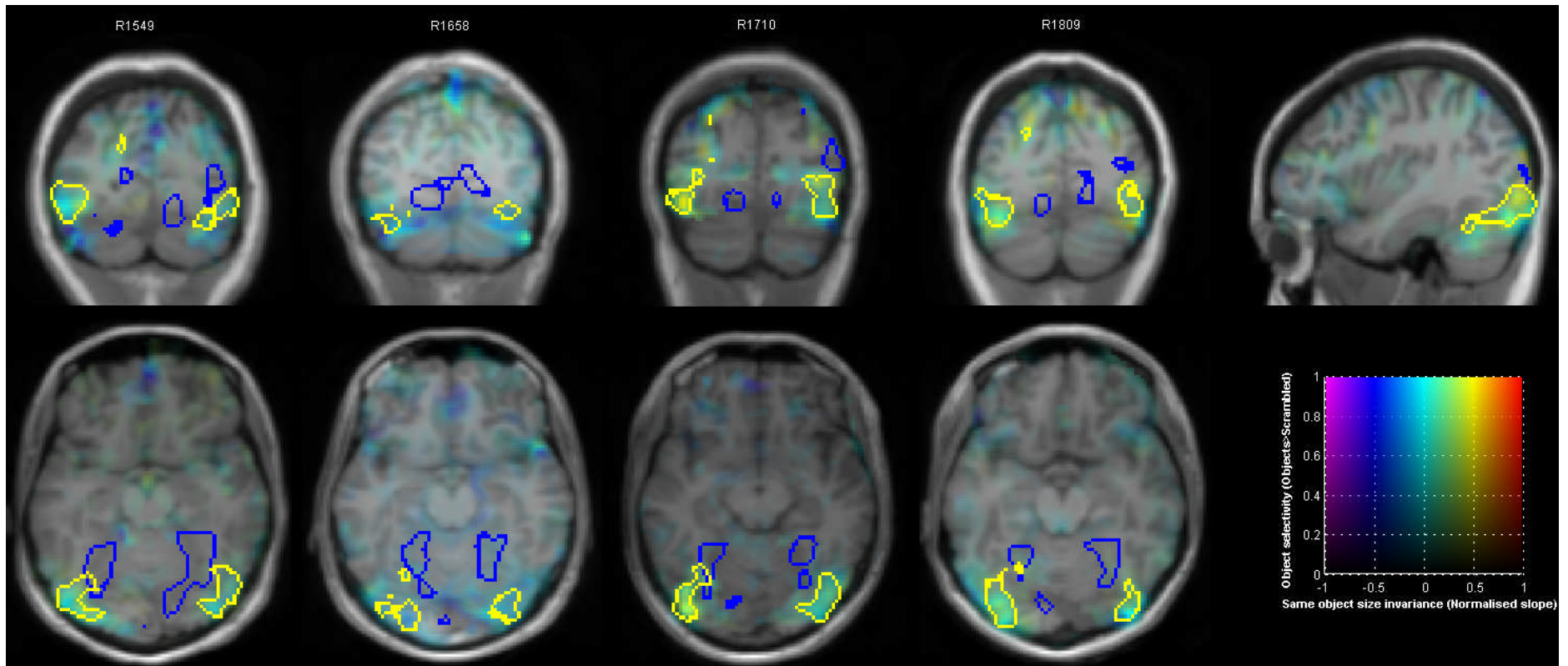
Colour determined by the angle formed between selectivity measures for objects (green) and places (blue), and opacity from the length of the vector formed by the two variables. Outlines of place (blue, places>faces) and object (yellow, objects>scrambled) selective regions defined in each individual using conventional localiser contrast and a minimum threshold of  $p < 0.001$  uncorrected. Slices in all figures are centred on MNI coordinates -25.63, -55.33, -19.66 and are displayed using neurological convention with the right hemisphere displayed on the right.



**Figure D.2.** Colour maps of landscape view invariance in place selective cortex for four individual transformed to MNI space (FLIRT). View invariance colour scale defined by normalised slope parameter of linear fit across four adaptation view shift conditions from the experiment in chapter 2. Voxels coloured red denote high levels of sensitivity, whereas voxels coloured cyan indicate an invariant response profile. Place selectivity determined by effect size measure in Place>Scrambled comparison in functional localiser, and used to determine voxel opacity. Outlines of place (blue, places>faces) and object (yellow, objects>scrambled) selective regions defined in each individual using conventional localiser contrast and a minimum threshold of  $p < 0.001$  uncorrected. Slices in all figures are centred on MNI coordinates -26.90, -49.63, -13.96 and are displayed using neurological convention with the right hemisphere displayed on the right.



**Figure D.3.** Colour maps of object view invariance in object selective cortex for four individuals transformed to MNI space (FLIRT). View invariance colour scale defined by normalised slope parameter of linear fit across four adaptation view shift conditions from the experiment in chapter 3. Voxels coloured red denote high levels of sensitivity, whereas voxels coloured cyan indicate an invariant response profile. Object selectivity determined by effect size measure in Objects>Scrambled comparison in functional localiser, and used to determine voxel opacity. Outlines of place (blue, places>faces) and object (yellow, objects>scrambled) selective regions defined in each individual using conventional localiser contrast and a minimum threshold of  $p < 0.001$  uncorrected. Slices in all figures are centred on MNI coordinates  $-40.82$ ,  $-78.73$ ,  $-15.86$  and are displayed using neurological convention with the right hemisphere displayed on the right.



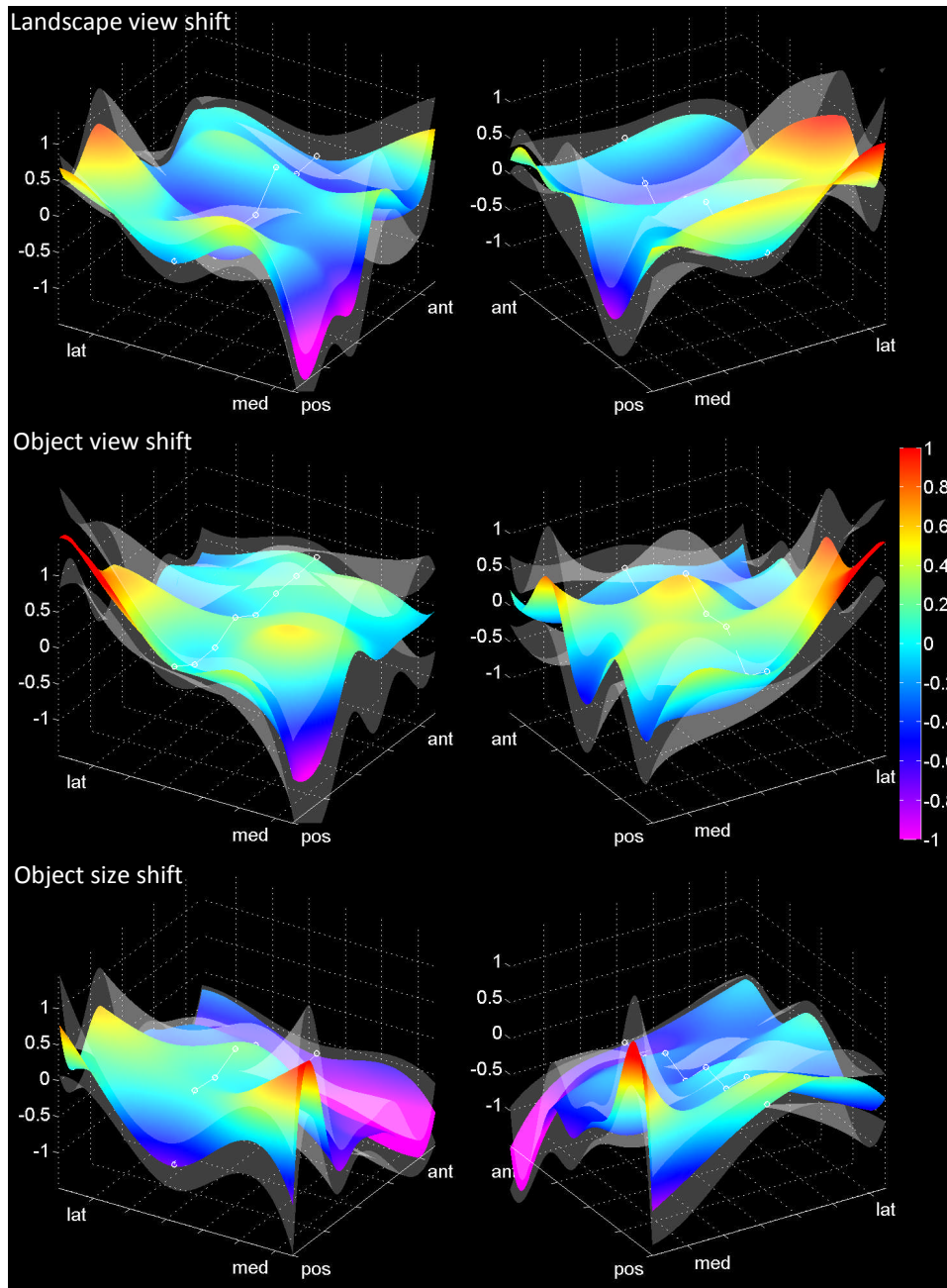
**Figure D.4.** Colour maps of same object size invariance in object selective cortex for four individuals transformed to MNI space (FLIRT). Size invariance colour scale defined by normalised slope parameter of linear fit across four adaptation size shift conditions (same objects) from the experiment in chapter 4. Voxels coloured red denote high levels of sensitivity, whereas voxels coloured cyan indicate an invariant response profile. Object selectivity determined by effect size measure in Objects>Scrambled comparison in functional localiser, and used to determine voxel opacity. Outlines of place (blue, places>faces) and object (yellow, objects>scrambled) selective regions defined in each individual using conventional localiser contrast and a minimum threshold of  $p < 0.001$  uncorrected. Slices in all figures are centred on MNI coordinates -38.29, -78.73, -16.49 and are displayed using neurological convention with the right hemisphere displayed on the right.

*Spatial analysis*

Statistical methods described in section **5.2.7** are used to compare the sensitivity patterns for same identity stimuli in all three experiments, and to assess the reliability of the patterns described in whole brain maps. Data extracted from the grid points in each individual were averaged together and displayed as surface plots for the left and right hemisphere separately for all three parametric adaptation experiments (**Figure D.2, D.2, & D.3**). These data from each of the experiments conducted are in standardised units (z-scores) and can be subject to within-subject ANOVA in order to determine the statistical reliability of the patterns observed. In order to compare the spatial pattern of invariance across the three experiments in this restricted sample of individuals, a repeated measures ANOVA was carried out, 3x2x5x8 repeated measures (experiment, hemisphere, medial-lateral dimension, posterior-anterior dimension). The patterns were entirely consistent across hemispheres, there were no main effects or interactions involving hemisphere, and as such hemisphere statistics are not reported. The surface data for each invariance measure are plotted in **Figure D.5**, and the statistical comparisons are shown in **Table D.1**.

**Table D.1.** *Same identity invariance spatial grid ANOVA*

<i>Experiment</i>	$F_{2,6} < 1, P=0.542$
<i>Med-lat dimension</i>	$F_{4,12} = 5.721, P < 0.01$
<i>Post-ant dimension</i>	$F_{7,21} = 3.796, P < 0.01$
<i>Experiment x med-lat</i>	$F_{8,24} = 0.258, P = 0.974$
<i>Experiment x post-ant</i>	$F_{14,42} = 2.065, P < 0.05$
<i>Post-ant x med-lat</i>	$F_{28,84} = 1.825, P < 0.05$
<i>Experiment x Post-ant x med-lat</i>	$F_{56,168} = 2.262, P < 0.001$



**Figure D.5.** *Landscape view-, object view-, and object size-invariance surface plots. View invariance measures derived from the slope of linear fit across parametric same identity shift conditions; landscape view shift (upper row), object view shift (middle row), object size shift (lower row). All data extracted from anatomical MNI grid coordinates in z-score standardised units. Red colour denotes high levels of sensitivity, whereas cyan indicates an invariant response profile. Surfaces generated from restricted sample of four subjects who took part in all experiments. Data from left and right hemispheres are plotted separately on the left and right respectively. Surface between grid points generated using spline interpolation function. Error surfaces (white) above and below main surface display 95% confidence intervals. Collateral sulcus points marked in white along the central longitudinal axis of each surface.*



These results indicate that the spatial pattern of sensitivity throughout the ventral surface of the brain is reliably different across all three experiments. This is true along the posterior-anterior dimension, and across the whole grid. The lack of a main effect of experiment and the overall effects of the spatial dimensions suggest that observed differences are not due to overall differences between levels of activation. The spatial patterns are much more similar for the two view manipulations. The effect of experiment is likely to be largely driven by the object size shift data, which shows more variation in sensitive in the medial part of the grid.

*Discussion: the robustness of spatial maps of invariance in individuals.*

It is possible to show, both qualitatively and quantitatively, that the spatial patterns observed at the full group level, for both selectivity maps, and invariance maps, are observable in individual subjects. It is possible to detect reliable differences in the spatial profile across different forms of same identity invariance. In particular, apparent difference between the spatial profile of size invariance as compared to view invariance, is evident within this small subset of individuals. These results suggest that, using appropriate stimulus manipulations and measures of effect size, that the present mapping invariance approach can illustrate symmetry and structure even in very small numbers of subjects

## References

- Abbott, L. F., Chance, F. S., & Salinas, E. (2008). Gain Modulation: Applications and Mechanisms.
- Aggleton, J. P., & Brown, M. W. (1999). Episodic memory, amnesia, and the hippocampal-anterior thalamic axis. *The Behavioral and brain sciences*, *22*(3), 425–444.
- Aggleton, J. P., & Vann, S. D. (2004). Testing the importance of the retrosplenial navigation system: lesion size but not strain matters: a reply to Harker and Whishaw. *Neuroscience and Biobehavioral Reviews*, *28*(5), 525–531.
- Aguirre, G. K., & D'Esposito, M. (1999). Topographical disorientation: a synthesis and taxonomy. *Brain*, *122*(9), 1613–1628.
- Amaral, D. G., & Lavenex, P. (2007). Hippocampal neuroanatomy. *The hippocampus book*, *1*(3), 37–114.
- Andersen, R. A., & Mountcastle, V. B. (1983). The influence of the angle of gaze upon the excitability of the light-sensitive neurons of the posterior parietal cortex. *The Journal of Neuroscience*, *3*(3), 532–548.
- Andrews, T. J., Clarke, A., Pell, P., & Hartley, T. (2010). Selectivity for low-level features of objects in the human ventral stream. *NeuroImage*, *49*(1), 703–711.
- Ashbridge, E., Perrett, D. I., Oram, M. W., & Jellema, T. (2000). Effect of image orientation and size on object recognition: responses of single units in the macaque monkey temporal cortex. *Cognitive Neuropsychology*, *17*(1-3), 13–34.
- Bartfeld, E., & Grinvald, A. (1992). Relationships between orientation-preference pinwheels, cytochrome oxidase blobs, and ocular-dominance columns in primate striate cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(24), 11905–11909.
- Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, *94*(2), 115–147.
- Biederman, I., & Gerhardstein, P. C. (1993). Recognizing depth-rotated objects - evidence and conditions for 3-dimensional viewpoint invariance. *Journal of Experimental Psychology-Human Perception and Performance*, *19*(6), 1162–1182.
- Bird, C. M., & Burgess, N. (2008). The hippocampus and memory: insights from spatial processing. *Nature Reviews Neuroscience*, *9*(3), 182–194.

- Booth, M. C., & Rolls, E. T. (1998). View-invariant representations of familiar objects by neurons in the inferior temporal visual cortex. *Cerebral Cortex (New York, N.Y.: 1991)*, 8(6), 510–523.
- Brincat, S. L., & Connor, C. E. (2004). Underlying principles of visual shape selectivity in posterior inferotemporal cortex. *Nat Neurosci*, 7(8), 880–886.
- Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network: anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences*, 1124, 1–38.
- Bulthoff, H. H., Edelman, S. Y., & Tarr, M. J. (1995). How are 3-dimensional objects represented in the brain. *Cerebral Cortex*, 5(3), 247–260.
- Byrne, P., & Becker, S. (2007). Remembering the past and imagining the future. *Psychological review*, 114(2), 340–375.
- Cabeza, R., & Nyberg, L. (2000). Imaging Cognition II: An Empirical Review of 275 PET and fMRI Studies. *J. Cognitive Neuroscience*, 12(1), 1–47.
- Cohen, L., Jobert, A., Le Bihan, D., & Dehaene, S. (2004). Distinct unimodal and multimodal regions for word processing in the left temporal cortex. *NeuroImage*, 23(4), 1256–1270.
- Cohen, L., Lehéricy, S., Chochon, F., Lemer, C., Rivaud, S., & Dehaene, S. (2002). Language-specific tuning of visual cortex? Functional properties of the Visual Word Form Area. *Brain: A Journal of Neurology*, 125(Pt 5), 1054–1069.
- Connolly, A. C., Guntupalli, J. S., Gors, J., Hanke, M., Halchenko, Y. O., Wu, Y., et al. (2012). The Representation of Biological Classes in the Human Brain. *The Journal of Neuroscience*, 32(8), 2608–2618.
- Connor, C. E., Gallant, J. L., Preddie, D. C., & Van Essen, D. C. (1996). Responses in area V4 depend on the spatial relationship between stimulus and attention. *Journal of neurophysiology*, 75(3), 1306–1308.
- Connor, C. E., Preddie, D. C., Gallant, J. L., & Van Essen, D. C. (1997). Spatial Attention Effects in Macaque Area V4. *The Journal of Neuroscience*, 17(9), 3201–3214.
- Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical Surface-Based Analysis: I. Segmentation and Surface Reconstruction. *NeuroImage*, 9(2), 179–194.
- Desimone, R., Albright, T. D., Gross, C. G., & Bruce, C. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *J. Neurosci.*, 4(8), 2051–2062.
- Dicarlo, J. J., & Cox, D. (2007). Untangling invariant object recognition. *Trends in Cognitive Sciences*, 11(8), 333–341.
- Dobbins, A. C., Jeo, R. M., Fiser, J., & Allman, J. M. (1998). Distance Modulation of Neural Activity in the Visual Cortex. *Science*, 281(5376), 552–555.

- Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A Cortical Area Selective for Visual Processing of the Human Body. *Science*, *293*(5539), 2470–2473.
- Dumoulin, S. O., Bittar, R. G., Kabani, N. J., Baker, C. I., Le Goualher, G., Pike, G. B., et al. (2000). A New Anatomical Landmark for Reliable Identification of Human Area V5/MT: a Quantitative Analysis of Sulcal Patterning. *Cereb. Cortex*, *10*(5), 454–463.
- Ekstrom, A. D., Kahana, M. J., Caplan, J. B., Fields, T. A., Isham, E. A., Newman, E. L., et al. (2003). Cellular networks underlying human spatial navigation. *Nature*, *425*(6954), 184–187.
- Engel, S. A., Rumelhart, D. E., Wandell, B. A., Lee, A. T., Glover, G. H., Chichilnisky, E. J., et al. (1994). fMRI of human visual-cortex. *Nature*, *369*(6481), 525–525.
- Epstein, R. A. (2005). The cortical basis of visual scene processing. *Visual Cognition*, *12*(6), 954.
- Epstein, R. A. (2008). Parahippocampal and retrosplenial contributions to human spatial navigation. *Trends in Cognitive Sciences*, *12*(10), 388–396.
- Epstein, R. A., Graham, K. S., & Downing, P. E. (2003). Viewpoint-Specific Scene Representations in Human Parahippocampal Cortex. *Neuron*, *37*(5), 865–876.
- Epstein, R. A., Harris, A., Stanley, D., & Kanwisher, N. (1999). The parahippocampal place area: recognition, navigation, or encoding? *Neuron*, *23*(1), 115–125.
- Epstein, R. A., & Higgins, J. S. (2007). Differential parahippocampal and retrosplenial involvement in three types of visual scene recognition. *Cerebral Cortex (New York, N.Y.: 1991)*, *17*(7), 1680–1693.
- Epstein, R. A., Higgins, J. S., Jablonski, K., & Feiler, A. M. (2007). Visual Scene Processing in Familiar and Unfamiliar Environments. *J Neurophysiol*, *97*(5), 3670–3683.
- Epstein, R. A., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, *392*(6676), 598–601.
- Epstein, R. A., Parker, W. E., & Feiler, A. M. (2007). Where Am I Now? Distinct Roles for Parahippocampal and Retrosplenial Cortices in Place Recognition. *J. Neurosci.*, *27*(23), 6141–6149.
- Van Essen, D. C., & Maunsell, J. H. R. (1983). Hierarchical organization and functional streams in the visual cortex. *Trends in Neurosciences*, *6*(0), 370–375.
- Ewbank, M. P., Schluppeck, D., & Andrews, T. J. (2005). fMR-adaptation reveals a distributed representation of inanimate objects and places in human visual cortex. *Neuroimage*, *28*(1), 268–279.

- Fang, F., Murray, S. O., Kersten, D., & He, S. (2005). Orientation-tuned fMRI adaptation in human visual cortex. *Journal of Neurophysiology*, *94*(6), 4188–4195.
- Feinberg, T. E., Schindler, R. J., Ochoa, E., Kwan, P. C., & Farah, M. J. (1994). Associative visual agnosia and alexia without prosopagnosia. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *30*(3), 395–411.
- Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex (New York, N.Y.: 1991)*, *1*(1), 1–47.
- Fischl, B., Sereno, M. I., & Dale, A. M. (1999). Cortical Surface-Based Analysis: II: Inflation, Flattening, and a Surface-Based Coordinate System. *NeuroImage*, *9*(2), 195–207.
- Fisher, R. A. (1956). *Statistical methods and scientific inference* (Vol. viii). Oxford, England: Hafner Publishing Co.
- Friston, K. J., Rotshtein, P., Geng, J. J., Sterzer, P., & Henson, R. N. (2006). A critique of functional localisers. *NeuroImage*, *30*(4), 1077–1087.
- Gallant, J. L., Braun, J., & Van Essen, D. C. (1993). Selectivity for polar, hyperbolic, and Cartesian gratings in macaque visual cortex. *Science (New York, N.Y.)*, *259*(5091), 100–103.
- Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience*, *3*(2), 191–197.
- Gibson, J. J., & Radner, M. (1937). Adaptation, after-effect and contrast in the perception of tilted lines. I. Quantitative studies. *Journal of Experimental Psychology*, *20*(5), 453–467.
- Goebel, R., Khorram-Sefat, D., Muckli, L., Hacker, H., & Singer, W. (1998). The constructive nature of vision: direct evidence from functional magnetic resonance imaging studies of apparent motion and motion imagery. *European Journal of Neuroscience*, *10*(5), 1563–1573.
- Goesaert, E., & Op de Beeck, H. P. (2010). Continuous mapping of the cortical object vision pathway using traveling waves in object space. *NeuroImage*, *49*(4), 3248–3256.
- Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, *15*(1), 20–25.
- Grill-Spector, K., Henson, R. N., & Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, *10*(1), 14–23.
- Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research*, *41*(10-11), 1409–1422.

- Grill-Spector, K., Kushnir, T., Edelman, S., Avidan, G., Itzhak, Y., & Malach, R. (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron*, *24*(1), 187–203.
- Grill-Spector, K., Kushnir, T., Hendler, T., & Malach, R. (2000). The dynamics of object-selective activation correlate with recognition performance in humans. *Nature Neuroscience*, *3*(8), 837–843.
- Grill-Spector, K., & Malach, R. (2001). fMR-adaptation: a tool for studying the functional properties of human cortical neurons. *Acta Psychologica*, *107*(1-3), 293–321.
- Grill-Spector, K., & Malach, R. (2004). The human visual cortex. *Annu. Rev. Neurosci.*, *27*, 649–677.
- Gross, C. G., Rocha-Miranda, C. E., & Bender, D. B. (1972). Visual properties of neurons in inferotemporal cortex of the Macaque. *Journal of Neurophysiology*, *35*(1), 96–111.
- Haak, K. V., Renken, R., & Cornelissen, F. W. (2010). One Cortical Network for the Visual Perception of Scenes and Textures. *Journal of Vision*, *10*(7), 1226–1226.
- Habib, M., & Sirigu, A. (1987). Pure topographical disorientation: a definition and anatomical basis. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *23*(1), 73–85.
- Hafting, T., Fyhn, M., Molden, S., Moser, M. B., & Moser, E. I. (2005). Microstructure of a spatial map in the entorhinal cortex. *Nature*, *436*(7052), 801–806.
- Hanazawa, A., Komatsu, H., & Murakami, I. (2000). Neural selectivity for hue and saturation of colour in the primary visual cortex of the monkey. *European Journal of Neuroscience*, *12*(5), 1753–1763.
- Harris, A. (1980). *Visual Coding and Adaptability* (1st ed.). Psychology Press.
- Hartley, T., Bird, C. M., Chan, D., Cipolotti, L., Husain, M., Vargha-Khadem, F., et al. (2007). The hippocampus is required for short-term topographical memory in humans. *Hippocampus*, *17*(1), 34–48.
- Hassabis, D., Chu, C., Rees, G., Weiskopf, N., Molyneux, P. D., & Maguire, E. A. (2009). Decoding Neuronal Ensembles in the Human Hippocampus. *Current Biology*, *19*(7), 546–554.
- Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, *293*(5539), 2425–2430.
- Haxby, J. V., Guntupalli, J. S., Connolly, A. C., Halchenko, Y. O., Conroy, B. R., Gobbini, M. I., et al. (2011). A Common, High-Dimensional Model of the Representational Space in Human Ventral Temporal Cortex. *Neuron*, *72*(2), 404–416.

- Hecaen, H., & Angelergues, R. (1962). Agnosia for Faces (Prosopagnosia). *Arch Neurol*, 7(2), 92–100.
- Holmes, G. (1945). Ferrier Lecture: The Organization of the Visual Cortex in Man. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 132(869), 348–361.
- Horton, J. C., & Hoyt, W. F. (1991). The representation of the visual-field in human striate cortex - a revision of the classic holmes map. *Archives of Ophthalmology*, 109(6), 816–824.
- Hubel, D. H., Wiesel, T. N., & Levay, S. (1977). Plasticity of ocular dominance columns in monkey striate cortex. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, 278(961), 377–&.
- Humphrey, N., & Weiskrantz, L. (1969). Size constancy in monkeys with inferotemporal lesions. *The Quarterly journal of experimental psychology*, 21(3), 225–238.
- Hung, C. P., Kreiman, G., Poggio, T., & DiCarlo, J. J. (2005). Fast readout of object identity from macaque inferior temporal cortex. *Science*, 310(5749), 863–866.
- Ison, M. J., & Quiroga, R. Q. (2008). Selectivity and invariance for visual object perception. *Frontiers in Bioscience: A Journal and Virtual Library*, 13, 4889–4903.
- Ito, M., Tamura, H., Fujita, I., & Tanaka, K. (1995). Size and position invariance of neuronal responses in monkey inferotemporal cortex. *J Neurophysiol*, 73(1), 218–226.
- Janzen, G., & Turennout, M. van. (2004). Selective neural representation of objects relevant for navigation. *Nature Neuroscience*, 7(6), 673–677.
- Janzen, G., Wagensveld, B., & Turennout, M. van. (2007). Neural Representation of Navigational Relevance Is Rapidly Induced and Long Lasting. *Cerebral Cortex*, 17(4), 975–981.
- Jenkinson, M., Bannister, P., Brady, T. J., & Smith. (2002). Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. *NeuroImage*, 17(2), 825–841.
- Jernigan, T. L., Gamst, A. C., Fennema-Notestine, C., & Ostergaard, A. L. (2003). More 'mapping' in brain mapping: statistical comparison of effects. *Human Brain Mapping*, 19(2), 90–95.
- Johnson, E. N., Hawken, M. J., & Shapley, R. (2008). The Orientation Selectivity of Color-Responsive Neurons in Macaque V1. *The Journal of Neuroscience*, 28(32), 8096–8106.

- Julian, J. B., Fedorenko, E., Webster, J., & Kanwisher, N. (2012). An algorithmic method for functionally defining regions of interest in the ventral visual pathway. *NeuroImage*, *60*(4), 2357–2364.
- Kamitani, Y., & Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nature Neuroscience*, *8*(5), 679–685.
- Kanwisher, N. (2000). Domain specificity in face perception. *Nature Neuroscience*, *3*(8), 759–763.
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The Fusiform Face Area: A Module in Human Extrastriate Cortex Specialized for Face Perception. *J. Neurosci.*, *17*(11), 4302–4311.
- Klein, A., Andersson, J., Ardekani, B. A., Ashburner, J., Avants, B., Chiang, M.-C., et al. (2009). Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *NeuroImage*, *46*(3), 786–802.
- Kobatake, E., & Tanaka, K. (1994). Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *Journal of Neurophysiology*, *71*(3), 856–867.
- Kohn, A., & Movshon, J. A. (2004). Adaptation changes the direction tuning of macaque MT neurons. *Nature Neuroscience*, *7*(7), 764–772.
- Kourtzi, Z., Erb, M., Grodd, W., & Bulthoff, H. H. (2003). Representation of the Perceived 3-D Object Shape in the Human Lateral Occipital Complex. *Cereb. Cortex*, *13*(9), 911–920.
- Kourtzi, Z., & Grill-Spector, K. (2005). fMRI adaptation: a tool for studying visual representations in the primate brain. *Fitting the mind to the world* (pp. 173–188). Oxford University Press.
- Kourtzi, Z., & Kanwisher, N. (2000). Cortical regions involved in perceiving object shape. *Journal of Neuroscience*, *20*(9), 3310–3318.
- Krekelberg, B., Boynton, G. M., & van Wezel, R. J. A. (2006). Adaptation: from single cells to BOLD signals. *Trends in Neurosciences*, *29*(5), 250–256.
- Kriegeskorte, N., Formisano, E., Sorger, B., & Goebel, R. (2007). Individual faces elicit distinct response patterns in human anterior temporal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(51), 20600–20605.
- Landis, T., Cummings, J. L., Benson, D. F., & Palmer, E. P. (1986). Loss of topographic familiarity. An environmental agnosia. *Archives of Neurology*, *43*(2), 132–136.
- Larsson, J., & Heeger, D. J. (2006). Two Retinotopic Visual Areas in Human Lateral Occipital Cortex. *The Journal of Neuroscience*, *26*(51), 13128–13142.



- Lever, C., Wills, T., Cacucci, F., Burgess, N., & O'Keefe, J. (2002). Long-term plasticity in hippocampal place-cell representation of environmental geometry. *Nature*, *416*(6876), 90–94.
- Levy, I., Hasson, U., Avidan, G., Hendler, T., & Malach, R. (2001). Center-periphery organization of human object areas. *Nature Neuroscience*, *4*(5), 533–539.
- Lieberman, M. D., & Cunningham, W. A. (2009). Type I and Type II error concerns in fMRI research: re-balancing the scale. *Social Cognitive and Affective Neuroscience*, *4*(4), 423–428.
- Logothetis, N. K., Pauls, J., Bülthoff, H. H., & Poggio, T. (1994). View-dependent object recognition by monkeys. *Current Biology*, *4*(5), 401–414.
- Lueschow, A., Miller, E. K., & Desimone, R. (1994). Inferior Temporal Mechanisms for Invariant Object Recognition. *Cerebral Cortex*, *4*(5), 523–531.
- Maguire, E. A. (2001). The retrosplenial contribution to human navigation: A review of lesion and neuroimaging findings. *Scandinavian Journal of Psychology*, *42*(3), 225–238.
- Malach, R., Levy, I., & Hasson, U. (2002). The topography of high-order human object areas. *Trends in Cognitive Sciences*, *6*(4), 176–184.
- Malach, R., Reppas, J. B., Benson, D. F., Kwong, K. K., Jiang, H., Kennedy, W. A., et al. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(18), 8135–8139.
- Marr, D. (1983). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. W. H. Freeman.
- Marr, D., & Nishihara, H. K. (1978). Representation and Recognition of the Spatial Organization of Three-Dimensional Shapes. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, *200*(1140), 269–294.
- McKone, E., & Robbins, R. (2007). The evidence rejects the expertise hypothesis: Reply to Gauthier & Bukach. *Cognition*, *103*(2), 331–336.
- McNeil, J. E., & Warrington, E. K. (1993). Prosopagnosia: a face-specific disorder. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *46*(1), 1–10.
- Milner, A. D., & Goodale, M. A. (1995). *The visual brain in action*. Oxford University Press.
- Miyashita, Y., & Chang, H. S. (1988). Neuronal correlate of pictorial short-term memory in the primate temporal cortex Yasushi Miyashita. *Nature*, *331*, 68–70.
- Muller, R., & Kubie, J. (1987). The effects of changes in the environment on the spatial firing of hippocampal complex-spike cells. *J. Neurosci.*, *7*(7), 1951–1968.

- Nasr, S., Liu, N., Devaney, K. J., Yue, X., Rajimehr, R., Ungerleider, L. G., et al. (2011). Scene-Selective Cortical Regions in Human and Nonhuman Primates. *The Journal of Neuroscience*, *31*(39), 13771–13785.
- Nishimoto, S., Vu, A. T., Naselaris, T., Benjamini, Y., Yu, B., & Gallant, J. L. (2011). Reconstructing Visual Experiences from Brain Activity Evoked by Natural Movies. *Current Biology*, *21*(19), 1641–1646.
- Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*, *10*(9), 424–430.
- O’Keefe, J., & Conway, D. H. (1978). Hippocampal place units in the freely moving rat: Why they fire where they fire. *Experimental Brain Research*, *31*(4).
- O’Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. Clarendon Press.
- Op de Beeck, H. P., Haushofer, J., & Kanwisher, N. (2008). Interpreting fMRI data: maps, modules and dimensions. *Nat Rev Neurosci*, *9*(2), 123–135.
- Park, S., & Chun, M. M. (2009). Different roles of the parahippocampal place area (PPA) and retrosplenial cortex (RSC) in panoramic scene perception. *NeuroImage*, *47*(4), 1747–1756.
- Pasupathy, A., & Connor, C. E. (1999). Responses to Contour Features in Macaque Area V4. *J Neurophysiol*, *82*(5), 2490–2502.
- Peelen, M. V., & Downing, P. E. (2005). Selectivity for the Human Body in the Fusiform Gyrus. *J Neurophysiol*, *93*(1), 603–608.
- Peng, X., & Van Essen, D. C. (2005). Peaked Encoding of Relative Luminance in Macaque Areas V1 and V2. *Journal of Neurophysiology*, *93*(3), 1620–1632.
- Pouget, A., & Sejnowski, T. J. (1997). A new view of hemineglect based on the response properties of parietal neurones. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *352*(1360), 1449–1459.
- Presson, C. C., DeLange, N., & Hazelrigg, M. D. (1989). Orientation specificity in spatial memory: What makes a path different from a map of the path? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(5), 887–897.
- Reinke, K., Fernandes, M., Schwindt, G., O’Craven, K., & Grady, C. L. (2008). Functional specificity of the visual word form area: general activation for words and symbols but specific network activation for words. *Brain and Language*, *104*(2), 180–189.
- Riesenhuber, M., & Poggio, T. (2000). Models of object recognition. *Nature Neuroscience*, *3 Suppl*, 1199–1204.
- Rolls, E. T. (1994). Brain mechanisms for invariant visual recognition and learning. *Behavioural Processes*, *33*(1–2), 113–138.

- Rudge, P., & Warrington, E. K. (1991). Selective impairment of memory and visual perception in splenial tumours. *Brain*, *114*(1), 349–360.
- Rust, N. C., & DiCarlo, J. J. (2010). Selectivity and Tolerance ('Invariance') Both Increase as Visual Information Propagates from Cortical Area V4 to IT. *J. Neurosci.*, *30*(39), 12978–12995.
- Salinas, E., & Abbott, L. F. (1997). Invariant visual responses from attentional gain fields, *77*.
- Salinas, E., & Thier, P. (2000). Gain modulation: A major computational principle of the central nervous system. *Neuron*, *27*(1), 15–21.
- Sawamura, H., Georgieva, S., Vogels, R., Vanduffel, W., & Orban, G. A. (2005). Using Functional Magnetic Resonance Imaging to Assess Adaptation and Size Invariance of Shape Processing by Humans and Monkeys. *J. Neurosci.*, *25*(17), 4294–4306.
- Sawamura, H., Orban, G. A., & Vogels, R. (2006). Selectivity of Neuronal Adaptation Does Not Match Response Selectivity: A Single-Cell Study of the fMRI Adaptation Paradigm. *Neuron*, *49*(2), 307–318.
- Saxe, R., Brett, M., & Kanwisher, N. (2006). Divide and conquer: A defense of functional localizers. *NeuroImage*, *30*(4), 1088–1096.
- Schwartz, E. L., Desimone, R., Albright, T. D., & Gross, C. G. (1983). Shape Recognition and Inferior Temporal Neurons. *Proceedings of the National Academy of Sciences of the United States of America*, *80*(18), 5776–5778.
- Schwarzlose, R. F., Baker, C. I., & Kanwisher, N. (2005). Separate face and body selectivity on the fusiform gyrus. *Journal of Neuroscience*, *25*(47), 11055–11059.
- Scoville, W. B., & Milner, A. D. (1957). Loss of Recent Memory After Bilateral Hippocampal Lesions. *Journal of Neurology, Neurosurgery & Psychiatry*, *20*(1), 11–21.
- Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., et al. (1995). Borders of multiple visual areas in humans revealed by functional magnetic-resonance-imaging. *Science*, *268*(5212), 889–893.
- Sereno, M. I., Pitzalis, S., & Martinez, A. (2001). Mapping of contralateral space in retinotopic coordinates by a parietal cortical area in humans. *Science*, *294*(5545), 1350–1354.
- Sheinberg, D. L., & Logothetis, N. K. (1997). The Role of Temporal Cortical Areas in Perceptual Organization. *Proceedings of the National Academy of Sciences*, *94*(7), 3408–3413.

- Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, *23 Suppl 1*, S208–219.
- Spiridon, M., Fischl, B., & Kanwisher, N. (2006). Location and spatial profile of category-specific regions in human extrastriate cortex. *Human Brain Mapping*, *27*(1), 77–89.
- Spiridon, M., & Kanwisher, N. (2002). How distributed is visual category information in human occipito-temporal cortex? An fMRI study. *Neuron*, *35*(6), 1157–1165.
- Squire, L. R., Stark, C. E. L., & Clark, R. E. (2004). The medial temporal lobe. *Annual Review of Neuroscience*, *27*, 279–306.
- Starrfelt, R., & Gerlach, C. (2007). The visual what for area: words and pictures in the left fusiform gyrus. *NeuroImage*, *35*(1), 334–342.
- Svoboda, E., McKinnon, M. C., & Levine, B. (2006). The functional neuroanatomy of autobiographical memory: A meta-analysis. *Neuropsychologia*, *44*(12), 2189–2208.
- Tanaka, K. (1996). Inferotemporal cortex and object vision. *Annual Review of Neuroscience*, *19*, 109–139.
- Tarr, M. J., & Bulthoff, H. H. (1998). Image-based object recognition in man, monkey and machine. *Cognition*, *67*(1-2), 1–20.
- Tarr, M. J., & Gauthier, I. (2000). FFA: a flexible fusiform area for subordinate-level visual processing automatized by expertise. *Nature Neuroscience*, *3*(8), 764–769.
- Tarr, M. J., & Pinker, S. (1989). Mental rotation and orientation-dependence in shape recognition. *Cognitive Psychology*, *21*(2), 233–282.
- Taylor, J. C., Wiggett, A. J., & Downing, P. E. (2007). Functional MRI analysis of body and body part representations in the extrastriate and fusiform body areas. *Journal of Neurophysiology*, *98*(3), 1626–1633.
- Tootell, R. B., Devaney, K. J., Young, J. C., Postelnicu, G., Rajimehr, R., & Ungerleider, L. G. (2008). fMRI mapping of a morphed continuum of 3D shapes within inferior temporal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(9), 3605–3609.
- Tootell, R. B., Reppas, J. B., Kwong, K. K., Malach, R., Born, R. T., Brady, T. J., et al. (1995). Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *The Journal of Neuroscience*, *15*(4), 3215–3230.
- Ullman, S. (1989). Aligning pictorial descriptions: An approach to object recognition. *Cognition*, *32*(3), 193–254.

- Ullman, S., & Basri, R. (1991). Recognition by linear combinations of models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *13*(10), 992–1006.
- Ullman, S., & Soloviev, S. (1999). Computation of pattern invariance in brain-like structures. *Neural Networks*, *12*(7–8), 1021–1036.
- Ungerleider, L. G., Ganz, L., & Pribram, K. H. (1977). Size constancy in rhesus monkeys: effects of pulvinar, prestriate, and inferotemporal lesions. *Experimental Brain Research. Experimentelle Hirnforschung. Expérimentation Cérébrale*, *27*(3-4), 251–269.
- Ungerleider, L. G., & Mishkin, M. (1982). Two Cortical Visual Systems. In M. A. Ingle, M. A. Goodale, & R. J. W. Mansfield (Eds.), *Analysis of Visual Behaviour*. MIT Press.
- Vann, S. D., Aggleton, J. P., & Maguire, E. A. (2009). What does the retrosplenial cortex do? *Nat Rev Neurosci*, *10*(11), 792–802.
- Vuilleumier, P., Henson, R. N., Driver, J., & Dolan, R. J. (2002). Multiple levels of visual object constancy revealed by event-related fMRI of repetition priming. *Nature Neuroscience*, *5*(5), 491–499.
- Wada, Y., & Yamamoto, T. (2001). Selective Impairment of Facial Recognition Due to a Haematoma Restricted to the Right Fusiform and Lateral Occipital Region. *Journal of Neurology, Neurosurgery & Psychiatry*, *71*(2), 254–257.
- Wandell, B. A., Chial, S., & Backus, B. T. (2000). Visualization and measurement of the cortical surface. *Journal of Cognitive Neuroscience*, *12*(5), 739–752.
- Wegman, J., & Janzen, G. (2011). Neural Encoding of Objects Relevant for Navigation and Resting State Correlations with Navigational Ability. *Journal of Cognitive Neuroscience*, 1–14.
- Weigelt, S., Kourtzi, Z., Kohler, A., Singer, W., & Muckli, L. (2007). The Cortical Representation of Objects Rotating in Depth. *The Journal of Neuroscience*, *27*(14), 3864–3874.
- Weigelt, S., Muckli, L., & Kohler, A. (2008). Functional magnetic resonance adaptation in visual neuroscience. *Reviews in the Neurosciences*, *19*(4-5), 363–380.
- Willenbockel, V., Sadr, J., Fiset, D., Horne, G., Gosselin, F., & Tanaka, K. (2010). The SHINE Toolbox for Controlling Low-Level Image Properties. *Journal of Vision*, *10*(7), 653–653.
- Wolbers, T., & Buchel, C. (2005). Dissociable Retrosplenial and Hippocampal Contributions to Successful Formation of Survey Representations. *J. Neurosci.*, *25*(13), 3333–3340.

Zipser, D., & Andersen, R. A. (1988). A back-propagation programmed network that simulates response properties of a subset of posterior parietal neurons. *Nature*, 331(6158), 679–684.