Pattern codes for perceived gaze direction revealed by functional MRI

Johan D. Carlin

Hughes Hall

University of Cambridge

This dissertation is submitted for the degree of

Doctor of Philosophy

Preface

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except where specifically indicated in the text.

This dissertation is within the word limit requirements set by the Biology Degree Committee and the Board of Graduate Studies.

To my mother, father and Elvis.

Acknowledgements

It is unlikely that this thesis would now be finished if a number of key people had not offered me help and support along the way. First I thank my supervisor, James Rowe, who provided essential technical advice and guidance and who remained enthusiastic about my work even when the research veered far from the Rowe lab's usual interests. I also thank my second supervisor, Andy Calder, who helped me develop theoretical and conceptual foundations and who consistently showed great attention to detail when critically examining my ideas (and who is also good at spotting extraneous commas). I am fortunate in that my two supervisors played complementary roles in providing a balance of technical and scientific expertise, and in offering encouragement and critical evaluation.

There are too many others to thank but I'll name a few. Raliza Stoyanova for moral support and for proof-reading (after 3 years, I've only just been told that it's 'full width at half *maximum*', not *mean*. Please don't check what I used in my papers), Noham Wolpe for submitting this dissertation in my absence (if you know Cambridge administration you realise this is no small effort). Russell Thompson and Niko Kriegeskorte for guiding me through the intricacies of the multivariate pattern methods upon which this research depends. Hamed Nili, Apoorva Bhandari, Sasha Vicente-Grabovetsky, Rhodri Cusack, Michael Ewbank, Peter Watson, Olaf Hauk and Ian Nimmo-Smith also shared their methodological insights at one point or another. All past and present members of my two supervisors' labs for fun times at social outings and for patiently suffering through many, many of my practice talks at lab meetings. Tristan Bekinschtein, Max Garagnani and the other CBU coffee aficiandos for their unwavering commitment to decent coffee, even if this requires hiding behind a shed

in the garden with a Bunsen burner and a Bialetti espresso brewer. Davy Evans and Rick Chen for turning their heads for me (and all the others who made a brave effort but didn't make the cut. Next time!). My participants for agreeing to lie in the scanner watching Davy and Rick turn their heads for 80 minutes straight.

Doris Tsao and Ralph Adolphs are owed special thanks for supporting a visit to Caltech and for offering me the opportunity to analyse macaque functional magnetic resonance imaging (fMRI) data. Their generosity and willingness to consider my ideas made a deep impression on me.

Finally, I thank the Medical Research Council for funding my PhD and for supporting the inspiring intellectual environment that is the Cognition and Brain Sciences Unit.

Pattern codes for perceived gaze direction revealed by functional MRI

Johan D. Carlin

Perceiving the direction of another's attention is a critical component of normal social behaviour. Seminal electrophysiology studies demonstrated that single cells in macaque superior temporal sulcus (STS) are tuned to specific directions of social cues, including gaze direction, head view, and body posture. Furthermore, a subset of such neurons respond to a single direction across multiple cues, suggesting that the code is driven by the direction of another's social attention regardless of how this is conveyed.

Attempts to reveal similar gaze representations in humans using fMRI have provided mixed results. This thesis describes research where multivariate pattern analysis (MVPA) methods are applied to fMRI data in order to better explain how the human brain and particularly STS codes perceived gaze direction.

After describing the MVPA methods applied in this thesis, I first demonstrate that fMRI response patterns in anterior STS distinguish between the direction of dynamic head turns, but not between the direction of rotation in non-social ellipsoids. In subsequent work, anterior STS is found to code the direction of another's gaze in a head view-invariant manner, thus demonstrating a potential parallel to previous macaque evidence for single cells that code the direction of another's attention. However, comparisons that run both across species (macaque, human) and methods (electrophysiology, fMRI) are problematic. To overcome this limitation I next tested whether macaque STS distinguishes gaze direction and head view when responses are measured with fMRI.

In conclusion, this thesis demonstrates the utility of applying MVPA to fMRI data to reveal socially-relevant representations of the direction of another's attention. The thesis particularly highlights anterior STS as a key region in supporting direction-specific representations of social cues. These results advance our understanding of how the brain codes socially-relevant information, and highlight possible similarities and dissimilarities between humans and macaques.

Contents

List of Figures

List of Tables

List of Acronyms

AL anterior lateral face patch

ANOVA analysis of variance

BOLD blood oxygen-level dependent

eCoG electrocorticography

EPI echoplanar imaging

FDR false discovery rate

FFA fusiform face area

fMRI functional magnetic resonance imaging

FSL FMRIB software library

FWE familywise error

FWHM full width at half maximum

GLM generalised linear model

IT inferotemporal cortex

LFP local field potentials

LIP lateral intraparietal area

MEG magnetoencephalography

MF middle fundus face patch

MION monocrystalline iron oxide nanoparticle

ML middle lateral face patch

MNI Montreal Neurological Institute

MPRAGE magnetisation prepared rapid gradient echo

MRI magnetic resonance imaging

MT middle temporal area

MTG middle temporal gyrus

MUA multiunit activity

MVPA multivariate pattern analysis

PET positron emission tomography

PL posterior lateral face patch

PyMVPA multivariate pattern analysis in python

RDM representational dissimilarity matrix

ROI region of interest

RSA representational similarity analysis

SnPM statistical non-parametric mapping

SPM statistical parametric mapping

STG superior temporal gyrus

STS superior temporal sulcus

SVM support vector machine

TE time to echo

TR time to repeat

Chapter 1

Introduction

Humans, macaques and many other animals have eyes that are arranged so that maximal visual acuity is achieved at the point of fixation (the fovea). This arrangement makes it possible for one individual to see where another is looking by following the direction implied by their iris. Such gaze perception skills are likely to play an important role in many aspects of social cognition but their precise neural underpinnings remain poorly understood.

In this thesis, I examine how the brain represents visual information about the direction of another's attention whether conveyed by eye gaze, dynamic head turns or static head view. I combine functional magnetic resonance imaging (fMRI) with multivariate pattern analysis (MVPA) to ask a number of fundamental questions about the neural mechanisms of gaze perception and other social cues. Which brain regions distinguish the direction of another's attention? How do such coding schemes operate? Are gaze direction codes similar in humans and in macaques? I begin by reviewing the existing literature on gaze perception before outlining my experimental approach to these questions at the end of the chapter.

1.1 The social relevance of another's gaze

The assertion that gaze perception is important to social cognition can be found in many of the papers reviewed in this chapter. Yet, what is the evidence that gaze really plays a functional role in social behaviour?

1.1.1 Gaze following and gaze cueing

Human infants begin to look where others are looking as early as 2 months of age (Scaife & Bruner, 1975), which suggests an important role for gaze following in development. In adults, a large literature has studied the attentional cueing effects of perceived gaze (for a review, see Frischen, Bayliss & Tipper, 2007). The typical finding is that following a brief presentation of a gazing face, participants are faster to respond to targets that appear in a location that is congruent with the gaze direction, even though the participants are instructed that the gaze cue is non-predictive of the target location (Friesen & Kingstone, 1998) or even counter-predictive of it (Driver, Davis, Kidd, Maxwell & Baron-Cohen, 1999). These results suggest that much like other cues that are spatially compatible with the location of a target, such as arrows (Tipples, 2002, 2008) or even laterally extended tongues (Downing, Dodds & Bray, 2004), averted gaze evokes a reflexive attentional reorienting response in the gazed-at direction.

Gaze cueing effects are generally interpreted in terms of covert attention, but few gaze cueing experiments have monitored fixation performance. Indeed, when eye tracking data was collected during cueing tasks, the gaze stimuli evoked small eye movements in the cued direction (Mansfield, Farroni & Johnson, 2003; Deaner & Platt, 2003), much in the same way that non-social peripheral cues evoke micro-saccades in the cued direction (Engbert & Kliegl, 2003). Gaze following, whether covert or overt, appears to be obligatory.

1.1.2 Eye contact effects

A fundamental distinction in gaze perception research is between averted and direct gaze. Eye contact slows reaction times to detecting subsequently presented lateral targets, relative to an upward (incongruent) averted gaze cue (Senju & Hasegawa, 2005). This suggests a special status or salience for direct gaze in attentional capture (Senju & Johnson, 2009). Infants' gaze following responses are elicited more reliably when averted gaze is preceded by eye contact (Senju & Csibra, 2008) so the exact sequence of direct and averted gaze may also play an important role in social behaviour.

The attentional capture effects of eye contact also appear in the 'stare in the crowd' effect, where participants are faster to detect a direct gaze face amongst averted gaze

distracter than they are to detect an averted gaze face amongst direct gaze distracters (Von Grüneau & Anston, 1995; Conty, Tijus, Hugueville, Coelho & George, 2006). Similarly, faces with direct gaze are detected more rapidly than faces with averted gaze under conditions of perceptual suppression (Stein, Senju, Peelen & Sterzer, 2011). Thus, whereas averted gaze cues attention to the gazed-at location, eye contact draws attention and awareness toward the gazing face itself (Senju & Johnson, 2009).

1.1.3 The role of gaze in social interaction

The literature I have discussed so far provides indirect evidence for the functional importance of gaze by showing that participants are responsive to perceived gaze direction even under conditions where gaze is irrelevant or even detrimental to the task at hand. The paradigms I turn to next demonstrate how gaze information is used during natural social interaction.

Gaze following is a useful strategy in many cooperative contexts. When pairs of participants are searching for the same target in an array, the target is detected more rapidly if the searchers have access to the other's point of gaze as implemented via an artificial gaze cursor on the screen (Brennan, Chen, Dickinson, Neider & Zelinsky, 2008). Similarly, during tasks where one participant verbally describes a target in an array while looking at it, the observing participant is able to identify the target before it has been unambiguously described (Hanna & Brennan, 2007). These findings suggest a role for gaze in guiding cooperative search.

Gaze perception also plays a role in disambiguation during language acquisition, where infants face the considerable challenge of learning which object a caregiver is talking about. Infants can learn such word-object mappings by following the speaker's gaze (Baldwin, 1991, 1993a,b). Indeed, the degree to which the mother refers to objects that the infant is gazing at correlates with the infant's vocabulary at 21 months of age, while the mother's references to objects the infant is not gazing at correlates negatively with vocabulary (Tomasello & Farrar, 1986). These findings highlight the interactive nature of gaze perception in communication. Not only does the infant follow the mother's gaze to disambiguate her verbal utterances but the mother also adapts her utterances to suit the infant's current focus of attention by following the infant's gaze.

Gaze also facilitates communication by mediating conversational turn-taking. A break in an utterance may indicate that the current speaker has finished their turn or alternatively that they are merely considering what to say next. Humans use gaze to disambiguate these cues. The speaker makes eye contact with the conversational partner if the break signals that a response is required, while averted gaze signals that the speaker wishes to hold the floor (Kendon, 1967). Thus, averted gaze can be used to convey other information than attentional shifts (cf. Section 1.1.1). Indeed, observers readily distinguish whether an actor's averted gaze conveys a shift in attention or floor-holding (Morency, Christoudias & Darrell, 2006). Importantly, this information is only available in dynamic videos; it does not appear to be present in single images of faces with averted gaze (Morency et al., 2006), suggesting that the dynamics of gaze shifts convey information about the social meaning of gaze.

Given its role in mediating social behaviour, it should come as no surprise that a person's gaze behaviour also influences how that person is evaluated on social dimensions. Participant likeability ratings are reduced both for individuals who consistently avert their gaze during social interaction and for individuals who never break eye contact (Argyle, Lefebvre & Cook, 1974). Individuals who maintain normal or high rates of eye contact are also rated as more credible and attractive during mock employment interviews (Burgoon, Manusov, Mineo & Hale, 1985). Thus, humans are sensitive not merely to the moment-to-moment direction of gaze and its social interpretation but also to the overall distribution of gaze shifts during the course of interactions with a particular individual.

1.1.4 Why study perception of gaze direction?

In this section I sought to demonstrate the functional relevance of gaze perception through two strands of research. First, cognitive and developmental experiments have demonstrated that another's gaze direction has a 'push-pull' effect on the observer, where averted gaze pushes covert attention or overt reorienting responses in the cued direction and direct gaze pulls attention toward the gazing individual. Second, more natural paradigms from developmental, social and applied psychology demonstrate that there are good reasons for humans to exhibit such sensitivity to gaze direction: this simple cue forms the basis for a range of sophisticated social functions includ-

ing cooperative search, language acquisition, conversational turn-taking and person perception. Such wide social relevance provides the fundamental justification for the focus on gaze direction representations in this thesis: all these social functions require a visual representation of another's gaze direction.

1.2 Gaze direction representations

The literature in this section concerns the psychophysics of discriminating gaze. The dependent measure is inconsistent in this literature, with some investigators describing gaze discrimination performance in terms of the physical displacement of the iris image on the observer's retina (referred to here as *iris eccentricity*), while others measure the shift produced on the actor's retina (*gaze angle*) or in terms of the absolute location of the target relative to the actor (*target eccentricity*). I used these concepts extensively in the following sections (for an illustration, see Figure 1.1).

1.2.1 Discriminating the target of another's gaze

Most gaze discrimination studies have focused on triadic tasks, where participants distinguish which of several targets is currently being looked at by an actor. Sensitivity is very high at such tasks with iris eccentricity thresholds of less than $1/120^\circ$ (Cline, 1967; Symons, Lee, Cedrone & Nishimura, 2004; Anstis, Mayhew & Morley, 1969) and gaze angle thresholds of around 1° at 1 m viewing distance (Symons et al., 2004). Such performance corresponds to the resolving power of normal visual acuity. Performance is generally better when gaze is directed toward targets close to the centre relative to targets further in the periphery (Symons et al., 2004; Cline, 1967) and for horizontal as opposed to vertically displaced targets (Cline, 1967).

1.2.2 Discriminating eye contact from averted gaze

In dyadic gaze tasks, participants are instead asked to distinguish between eye contact and averted gaze. An early study reported dyadic gaze discrimination thresholds of around $1/60^\circ$ iris eccentricity and a gaze angle of 2.8° at 2 m viewing distance (Gibson & Pick, 1963). Such performance is comparable to the previously discussed triadic

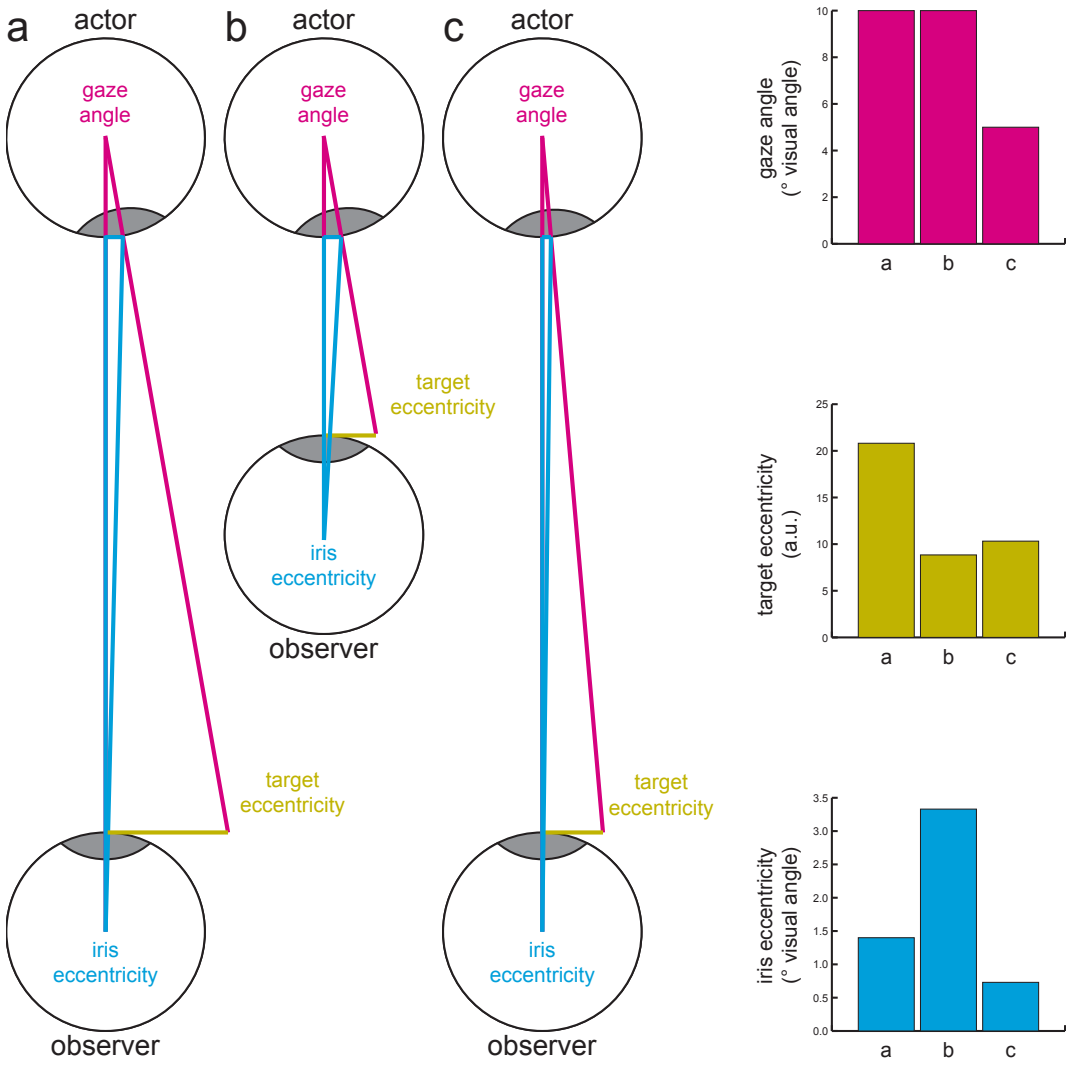


Figure 1.1: Schematic demonstration of how changes to gaze angle and viewing distance affect three different measures of gaze. (**a**) An actor (above) shifts his gaze rightward with a given gaze angle. This shift produces a displacement in iris eccentricity on the observer's retina (below) . If it is assumed that the new target of the actor's gaze is on the same plane as the observer's iris, the target eccentricity is the distance between the observer's iris and the new target in real-world coordinates. (**b**) If the viewing distance is decreased but the gaze angle is held constant, the target eccentricity is reduced, while the iris eccentricity is increased. (**c**) If the viewing distance is held constant but the gaze angle is reduced, target and iris eccentricity are both reduced. These effects are also shown in the bar charts in the far right panels.

gaze studies. To put these measures into context, consider that a gaze shift between the observer's left and right eye would constitute a gaze angle shift of approximately 1.8° at the same viewing distance, assuming that the observer had a normal inter-pupillary distance of 63 mm (Dodgson, 2004). Gibson & Pick (1963) concluded that participants judge any gaze direction up to the edge of the observer's face as conveying eye contact. Although this study indicated that thresholds for detecting eye contact are similar to thresholds for triadic gaze, these thresholds were estimated using the standard deviation of the mean, a method that may overestimate performance (Vine, 1971).

Subsequent studies that avoided this problematic threshold definition have instead argued that discrimination of dyadic gaze is coarser than discrimination of triadic gaze. Participants struggled to discriminate eye contact from gaze directed to other parts of the participant's face, such as the mouth (Lord & Haith, 1974), even though many of these points were at gaze angles that well exceeded the thresholds estimated by Gibson & Pick (1963). For instance, the mouth appeared at a gaze angle of 4.2° in the close viewing condition (103 cm viewing distance) but subjects rated gaze to this location as eye contact 46% of the time. Lord & Haith (1974) interpreted their data as evidence that gaze discrimination performance is poorer than previously thought. However, these two dyadic gaze studies actually agree that gaze is not reliably rated as averted until the target eccentricity extends past the observer's face (rather than merely away from the eyes). The contradictory thresholds for gaze angle appear to be a consequence of Lord & Haith (1974) sampling more points around the face at a closer distance, thus ensuring that stimuli with larger gaze angles and iris eccentricities would still fall on the face (Figure 1.1), while Gibson & Pick (1963) only used one target that fell on the observer's face and used a greater viewing distance. Collectively, both studies hint at a fundamental distinction between how participants code triadic and dyadic gaze: while triadic gaze is primarily limited by visual acuity, dyadic gaze thresholds are considerably coarser.

This notion was more closely examined by Gamer & Hecht (2007), who used a task where participants shifted the horizontal gaze direction of an averted gaze face until it was perceived to be gazing at them or a direct gaze face until it was perceived to look away. Results showed that participants rated a wide cone of gaze directions as signalling eye contact. In the experiment most comparable to the previously discussed

studies, the width of this cone was 3.7° gaze angle at 1 m viewing distance, which is comparable to the previous studies in that it roughly corresponds to the edge of the observer's face in target eccentricity (Gibson & Pick, 1963; Lord & Haith, 1974). Importantly, the gaze angle width of the cone narrowed with greater viewing distances, which means that participants were carrying out a finer discrimination of gaze angle and iris eccentricity at the farther viewing distance (Gamer & Hecht, 2007). Thus, the participants' judgements appeared to be based on the target eccentricity of their own face rather than iris eccentricity or gaze angle. When viewing a face from a close distance a wider range of gaze angles fall on the observer's face, while a longer distance means that a smaller range of gaze angles fall within the target eccentricity of the observer's face.

Across these dyadic gaze studies, gaze to the face is consistently rated as direct (Gamer & Hecht, 2007; Gibson & Pick, 1963) or is confusable with direct gaze (Lord & Haith, 1974). The width of this gaze cone is considerably larger than what would be expected by triadic gaze discrimination thresholds. This tendency to 'err on the side of caution' when it comes to potential eye contact may be motivated by the social importance of this particular gaze direction (Section 1.1.2). One useful way to think of the distinction between triadic and dyadic gaze performance is that triadic gaze thresholds are limited by the smallest detectable change in iris eccentricity independently of gaze angle and target eccentricity, while dyadic gaze thresholds are limited by the target eccentricity that moves gaze away from the observer's face independently of iris eccentricity and gaze angle (Figure 1.1).

1.2.3 Discriminating gaze direction across head view

The vast majority of gaze perception research relies on measuring the effects of perceived gaze in front-facing heads. Gaze perception is more challenging than this in the real world, because heads can be encountered in a range of views and the observer must thus infer that the same gaze direction can be conveyed by stimuli as distinct as direct gaze in a front-facing head and a turned head glancing sideways back at the observer.

Wollaston (1824) presented a classic illusion where superimposing the same eye region on drawings of two different head views produced a striking change in perceived

gaze direction (Figure 1.2). Subsequent studies that used similarly edited photos have demonstrated that the Wollaston illusion attracts gaze judgements towards the head view (Langton, Honeyman & Tessler, 2004; Todorović, 2006; Kluttz, Mayes, West & Kerby, 2009).

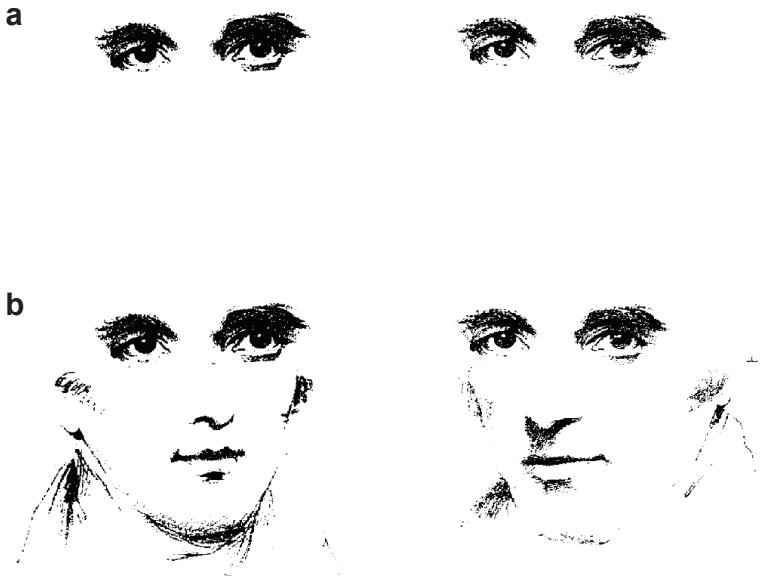


Figure 1.2: The Wollaston illusion demonstrates how identical eye regions (**a**) can produce different gaze percepts depending on the head outline on which they are superimposed (**b**). Adapted from Wollaston (1824).

The Wollaston illusion is frequently cited as evidence that averted head views bias perceived gaze away from the veridical direction (e.g., Langton & Bruce, 2000; Langton et al., 2004; Jenkins, 2007; Kluttz et al., 2009). However, it is important to note that the reason why the illusion works is because a real face with a head and eye region configuration matched to the Wollaston display would in fact be gazing roughly in the direction that participants perceive (Todorović, 2006). In this sense, there is no head view bias, but rather a reasonable interpretation of where a person with this particular head view and eye region configuration would be looking. Nevertheless, this work demonstrates that gaze direction judgements involve more than a simple view-based analysis of the eye region (I return to this notion in Section 1.2.5).

Studies that measured gaze discrimination performance across a range of non-artificial head views also observed that gaze judgements are biased by head view.

However, these effects are different from the Wollaston-based experiments in that the effects are less striking and importantly, the bias is generally in the opposite direction. Participants judge eye contact in a left-averted head as being directed to their right (Gibson & Pick, 1963) and similar repulsive head view effects also appear in triadic gaze paradigms (Cline, 1967; Anstis et al., 1969). This literature is not entirely consistent: Gamer & Hecht (2007) observed an attractive effect of head view on eye contact judgements, which is consistent with the Wollaston-based paradigms (Langton et al., 2004; Todorović, 2006; Kluttz et al., 2009). It is unclear why these results conflict but note that Gamer & Hecht (2007) used a smaller head view change (10°) than previous work (30° in Anstis et al., 1969; Cline, 1967; Gibson & Pick, 1963). It is thus possible that the effect of head view on gaze judgements is nonlinear but this has yet to be tested directly.

Gaze discrimination is not only biased by head view but is also less precise in the context of turned relative to front-facing heads. Increased gaze discrimination thresholds are observed in triadic tasks (Anstis et al., 1969; Cline, 1967) but not in dyadic tasks (Gibson & Pick, 1963; Gamer & Hecht, 2007). This pattern of effects provides further evidence for a special status for discrimination of eye contact (Section 1.1.2).

1.2.4 Contextual effects on gaze discrimination

A recurring theme in this chapter is that participants adopt somewhat different strategies when making triadic and dyadic gaze judgements. Triadic gaze codes are highly sensitive to subtle gaze deviations, while dyadic gaze codes involve a more generous criterion, where any gaze toward the face is interpreted as conveying eye contact (Section 1.2.2). However, the triadic gaze studies that obtained such strict thresholds generally used tightly-spaced arrays of targets (Symons et al., 2004) or continuous scales on which participants rated gaze (Anstis et al., 1969). Lobmaier, Fischer & Schwaninger (2006) asked subjects to judge gaze direction through mouse clicks (a continuous measure) but presented displays where an obvious target appeared in the vicinity of the gazed-at location. Participant's judgements of triadic gaze were systematically pulled in the direction of the target (Lobmaier et al., 2006). Thus, much like dyadic gaze judgements are pulled towards a particular salient object (the self), triadic gaze judgements can also be pulled in the direction of objects in the environment.

Dyadic gaze judgements are also biased by other factors. Participants judge a wider range of gaze directions as conveying eye contact in the context of an angry gazing face, relative to fearful and neutral faces (Ewbank, Jennings & Calder, 2009; Lobmaier, Tiddeman & Perrett, 2008), which may reflect the particular threat value of anger directed at the self (Ewbank et al., 2009). Happy expressions also elicit a greater degree of eye contact judgements than angry, fearful and neutral expressions (Lobmaier et al., 2008), which the authors interpreted in terms of a self-referential positivity bias. Indeed, dyadic gaze judgements can even be influenced by non-visual information: a greater range of gaze directions are judged to convey eye contact when the participant hears their own name relative to when they hear another's name (Stoyanova, Ewbank & Calder, 2010). These findings show that gaze perception is a fundamentally integrative process where judgements of another's gaze are influenced by contextual information not just from the face or even from the visual scene but also from other modalities.

1.2.5 Evaluating theories of gaze discrimination

The preceding sections have discussed a range of data on the characteristics of human gaze discrimination performance. The theoretical accounts of gaze discrimination that I turn to next are in a less developed state.

The simplest and most frequently cited class of gaze discrimination models posit that gaze direction is coded based on a view-based analysis of the eye region (Gibson & Pick, 1963; Symons et al., 2004; Anstis et al., 1969; Emery, 2000). For instance, gaze direction may be be approximated as the scleral ratio, that is, a simple size comparison of the visible sclera on either side of the iris (Ando, 2002; Anstis et al., 1969; Langton, Watt & Bruce, 2000).

Scleral ratio models are appealing because they posit a computationally simple mechanism that could plausibly be implemented by competitive interactions between visual neurons that respond to the two halves of the sclera (Ando, 2004). The model can explain a range of observations. For example, reducing the brightness of the sclera on one side biases gaze judgements in the manipulated direction (Ando, 2002, 2004). Similarly, contrast inversion of the eye region can induce an apparent reversal of gaze direction (Sinha, 2000; Ricciardelli, Baylis & Driver, 2000). This is easily explained since the scleral ratio will reverse together with contrast. The fundamental problem

with the scleral ratio model is that is fails when head view changes. For example, consider a person who maintains eye contact as he turns his head left to right. The scleral ratio will change even though the gaze direction is constant (Gibson & Pick, 1963; Todorović, 2006).

In principle, a visual analysis of the rotation of the eye suffices to discriminate gaze direction. Such iris pointing models (Kluttz et al., 2009; Todorović, 2006) posit a representation of the orientation of the eyeball relative to the observer. Unlike scleral ratio model, iris pointing models can explain how gaze is robustly estimated across different head views since the orientation of the eyeball is always toward the target. Instead, the iris pointing model suffers from the opposite problem in that is does not explain why gaze judgements are biased by head view (Section 1.2.3), if the eye is analysed in isolation (Todorović, 2006).

A natural solution to the shortcomings of the view-based eye region models is to include computation of head view. For example, one could posit multiple view-based representations where each combination of head view and eye region information is mapped to a particular gaze direction (Todorović, 2006). Alternatively gaze perception can viewed as a hierarchical process where head view trumps body posture if both are visible, while gaze direction trumps both head view and body posture whenever gaze information is present (Perrett & Emery, 1994). This direction-of-attention model assumes a form of iris pointing analysis, where the gaze direction can be encoded independently of head view and body posture. As such, it fails to predict interactions between gaze and head view (Langton et al., 2000).

In summary, there is no comprehensive theory of gaze discrimination that accounts for both how gaze is perceived across head views and how such performance can nevertheless be biased by head view. Few attempts have been made to consider contextual or eye contact influences on gaze judgements. There is a great need for a more formal, computational framework to explain gaze performance (I return to this point in Chapter 6).

1.3 Gaze direction after-effects

Perceptual after-effects are based on the principle that exposure to an adapting stimulus shifts subsequent perceptual judgements of a test stimulus away from the adaptor.

The studies in the preceding section used discrimination performance to infer underlying gaze direction representations. The adaptation studies that I discuss next aim to manipulate the content of such representations through adaptation in order to measure how this manipulation affects discrimination performance.

Initial adaptation studies focused on low level physical dimensions such as motion direction and orientation (Clifford, 2002). For example, adapting to downward drifting motion causes subsequent static test stimuli to apparently drift in the opposing, upward direction (Addams, 1834). This motion after-effect suggests the presence of opponent coding between representations coding upward and downward motion. Indeed, such opponent mechanisms can be observed directly in the retina of the rabbit when ganglion cells tuned to opposing motion directions are adapted (Barlow & Hill, 1963). This striking correspondence between a perceptual phenomenon and underlying physiological mechanisms highlights the power of the adaptation approach in revealing representational content.

Recently, similar adaptation effects have been reported for a number of high level face dimensions, including identity (Leopold, O'Toole, Vetter & Blanz, 2001), attractiveness (Rhodes, Jeffery, Watson, Clifford & Nakayama, 2003), expression, gender and ethnicity (Webster, Kaping, Mizokami & Duhamel, 2004). In this section I review research that uses adaptation to investigate directional representations for gaze and other social cues.

1.3.1 Adaptation to head view and gaze direction

The first indication that adaptation might be used to study gaze perception was from Fang & He (2005), who showed that adapting participants to left or right head views caused a shift in left-right judgements of subtly turned test heads relative to a baseline without adaptation. After adaptation to heads turned in one direction, heads turned in the same direction were more frequently rated as being turned in the opposing, non-adapted direction. By contrast, there was no effect of adaptation on test heads turned in the non-adapted direction. Similar view adaptation effects were observed for cars and wire frame objects. However, adapting to a turned wire frame did not shift judgements of head view stimuli, suggesting that separate view-specific representations were being adapted for each object category.

While head view adaptation is only subtly affected when the identity and/or gender changes between adaptor and test, inverting the adapting face disrupts adaptation to upright test faces (but does not abolish it, Fang & Ijichi, 2007). These findings show how adaptation can be used to investigate the nature of head view representations, which appear more tolerant to identity and/or gender changes than to inversion.

The interpretation of head view adaptation effects is limited by the fact that the adapting heads in the studies discussed so far included a visible eye region, which makes it difficult to determine whether the effects were driven by the head view or by the concurrently-changing gaze direction.

Studies that used averted gaze in front-facing heads as adaptors rather than head view also observed direction-specific adaptation effects (Seyama & Nagayama, 2006; Jenkins, Beaver & Calder, 2006). In light of the attentional cueing effects of eye gaze (Section 1.1.1) it is important to note that adapting to left and right headed arrows produces no gaze direction after-effects (Seyama & Nagayama, 2006), even though such arrows cue attention similarly to gaze (Tipples, 2008, 2002). Thus, the representation that is adapted appears to involve perceptual mechanisms, rather than attentional mechanisms such as inhibition of return (Klein, 2000).

1.3.2 Multi-channel coding

Adaptation to face dimensions such as identity is often interpreted in terms of opponent coding (Leopold et al., 2001; Rhodes & Jeffery, 2006), where a continuum is represented by the relative activity of two channels coding each extreme pole (for different implementations of this, see Webster & MacLeod, 2011). However, there is evidence that gaze direction and head view are instead coded in a multi-channel manner. Two key observations that support this: first, adapting to alternating left/right gaze directions makes participants more likely to rate subtly averted test stimuli as direct; second, adapting to direct gaze has the opposite effect, making participants more likely to rate test stimuli as averted (Calder, Jenkins, Cassel & Clifford, 2008). These opposing effects cannot by explained by two-channel or opponent-coded models, which predict that adapting to alternations between the two channels should adapt the channels similarly as adapting to the midpoint between them (that is, direct gaze in this case).

Similar evidence has been presented for head view (with closed eyes, unlike the

previously discussed studies, Lawson, Clifford & Calder, 2011) and body posture (Lawson, Clifford & Calder, 2009), suggesting that different cues to another's attention are coded according to similar principles. However, note that none of these findings necessarily imply the existence of any particular configuration of channels, such as a direct gaze channel (cf. Section 1.1.2). Instead, these findings are consistent with a large number of multi-channel coding schemes (Webster & MacLeod, 2011).

1.3.3 Adaptation across head and gaze

The preceding sections described how similar direction-specific adaptation effects have been observed for head view, body posture and gaze direction, which raises the question of whether the same underlying representation is being adapted in all cases.

A first suggestion that gaze adaptation may reflect head view-invariant representations came from Jenkins et al. (2006), who showed that gaze adaptation effects appear when the adapting faces have their heads turned in the direction of the gaze (cf. Fang & He, 2005) and the test faces gazed left, direct or right relative to a front-facing head. The adapted representation was thus invariant to the physical change between the adaptation and test stimuli. Bi, Su, Chen & Fang (2009) made a direct comparison between head view adaptation during congruent or incongruent gaze. Adaptation effects in head view test stimuli were reduced when the averted head view adaptors had their gaze turned back towards the observer, suggesting that adaptation to turned heads with open eyes is partially governed by the direction of gaze.

The direct test of whether adaptation to head, gaze and body reflect a shared representation would be to show that adapting to one of these cues leads to after-effects measurable when another cue type is used as the test stimulus. Teufel, Alexis, Todd, Lawrance-Owen, Clayton & Davis (2009) adapted participants to left or right turned heads where the eye region was obscured by opaque glasses. Adaptation transferred to gaze tests in a front-facing head but only when participants believed that the actor could see through the glasses. Such cross over between different social cues implies the presence of a single underlying representation of the direction of another's attention, although this representation appears to depend on how the actor's intentions are interpreted.

1.4 Neural representations of head view and gaze direction in macaques

Although the discussion thus far has centred on human perceptual performance, many other animals also exhibit behavioural sensitivity to another's gaze (for reviews, see Emery, 2000; Shepherd, 2010). In particular, rhesus macaques show a range of gaze effects, including gaze following (Ferrari, Kohler, Fogassi & Gallese, 2000) and gaze-cued micro-saccades and attentional shifts (Deaner & Platt, 2003, for human results see Section 1.1.1). This is useful, because as a consequence of the availability of invasive methods such as single unit recording, optical imaging and cortical ablation in macaques, the macaque visual system is arguably better-understood than the human system (Orban, Van Essen & Vanduffel, 2004), . These methods have revealed a wealth of information about how neural codes primarily in superior temporal sulcus (STS) represent the direction of another's attention.

Single neurons in STS are generally visually responsive, exhibit crossmodal tunings and are selective for certain directions of motion (Bruce, Desimone & Gross, 1981). This section concerns a subset of visually-selective cells that respond selectively to faces. Most studies reviewed here identified face-selective cells and restricted subsequent experimentation to this population. It is important to emphasise that face cells are only one group of STS cells, with estimates ranging from 2.7% (Bruce et al., 1981) to 76% (Jellema & Perrett, 2003) of visually-responsive cells in conventional electrophysiology, while estimates from fMRI-guided efforts are as high as 97% (Tsao, Freiwald, Tootell & Livingstone, 2006). Such wide discrepancies in the proportion of identified face cells may reflect differences in the exact STS recording site or the comprehensiveness of the tested face stimulus set. As I will discuss, many face cells in STS are highly selective for the intersection of a particular stimulus (such as heads) and a particular view (such as left profile). Cells with such specific tunings are easily missed if the test for face selectivity does not include faces with the appropriate configuration to drive the cell (for a related example, see Freiwald & Tsao, 2010).

1.4.1 Head view cells in macaque STS

The first evidence for face-selective cells was reported in macaque inferotemporal cortex (IT) (Gross, Rocha-Miranda & Bender, 1972) and evidence for face-selective cells in STS was first reported by Bruce et al. (1981). Subsequent work showed that face-selective STS cells are strongly selective for particular views (Perrett, Rolls & Caan, 1982) and in the following I refer to these as head view cells. The tuning of head view cells falls into two broad categories: some cells have unimodal tunings for a particular view (Perrett, Smith, Potter, Mistlin, Head, Milner & Jeeves, 1984). Most unimodal head view cells are tuned to views along the cardinal axes: front-facing, upward, downward, left or right profile (Perrett, Oram, Harries, Bevan, Hietanen, Benson & Thomas, 1991; Perrett et al., 1992). The dominant preferred view in this population is front-facing heads (Perrett et al., 1991). A second class of STS head view cells have bimodal tunings with similar responses to mirror-symmetric views, for instance to both left and right profiles (Perrett et al., 1984). More recent investigators have produced conflicting reports regarding how unimodal and bimodal head view cells are arranged in STS. One study found that unimodal cells are found anterior to bimodal cells (De Souza, Eifuku, Tamura, Nishijo & Ono, 2005), while a more recent study that used fMRI-guided electrophysiology found the reverse pattern (Freiwald & Tsao, 2010).

The cells discussed so far were responsive to static images of head views but another intriguing group of head view cells respond selectively to moving heads. Such cells may prefer head turns in a particular direction (Perrett, Smith, Mistlin, Chitty, Head, Potter, Broennimann, Milner & Jeeves, 1985; Hasselmo, Rolls, Baylis & Nalwa, 1989) or any motion that brings the head to face the observer (Perrett et al., 1985). The existence of the latter class of cells suggests that what is being represented is the endpoint of motion rather than the direction of motion as such. The presence of such cells suggests that dynamic information about social cues is represented differently from static information. However, the role of dynamic representations of head turns has not been explored in detail beyond these initial studies.

1.4.2 Gaze cells in macaque STS

The exact feature that drives head view cell tunings in these studies is somewhat unclear since the heads were generally presented with open eyes. It is thus equally plau-

sible that the tuning reflects gaze direction as head view (a related limitation appears in some of the behavioural adaptation studies, see Section 1.3.1). Surprisingly few macaque STS studies report gaze data or interactions between head view and gaze. One might speculate that this reflects a species difference: macaques and other animals with less visible sclera than humans may perceive the direction of another's attention more through head view than through gaze direction (Kobayashi & Kohshima, 1997) and by this logic one might expect neural representations of head view to be more prominent in macaques. However, note also that many face perception researchers have recorded STS neurons with the aim of identifying view-invariant face identity representations (for this interpretation of STS face cells, see e.g., Baylis, Rolls & Leonard, 1985; Tsao et al., 2008; Wang et al., 1998) and so had little reason to manipulate gaze direction.

When head view cells in anterior STS were tested for gaze responses, 64% were also responsive to gaze direction (Perrett, Smith, Potter, Mistlin, Head, Milner & Jeeves, 1985). For instance, a head view cell that preferred a profile view also preferred gaze in the same direction in a front-facing head, while a cell that preferred a frontal head view also preferred gaze toward the observer in a turned head. In all cases, responses to gaze were congruent with the head view preference. However, the degree of head view-invariance was not complete. Out of the 36 STS head view cells that were responsive to gaze direction, only 13 showed consistent preferences for a particular gaze direction across all tested head view and gaze configurations. The mechanisms underlying head view-invariant gaze direction codes may involve inhibition of conflicting head view information. Cells that preferred left gaze did not initially respond to a left head view when gaze was turned back toward the observer, but a response appeared when the eye region was obscured. Similarly, STS cells that are tuned to a particular body view also respond when an isolated head is presented in the same view (Wachsmuth, Oram & Perrett, 1994). Such tunings may thus reflect a joint representation of the direction of another's attention regardless of whether this is conveyed by head view, gaze direction or body posture (Perrett et al., 1992).

De Souza et al. (2005) also found STS head view cells with gaze direction responses. In their design, head view was presented in five horizontal increments and gaze was either congruent with the head view or directed back toward the observer. Head view cells in middle STS showed bimodal tunings to mirror-symmetric profile views and a minority showed gaze responses (45% exhibited gaze effects). Head view

cells in anterior STS showed unimodal tunings to one view and a majority showed gaze responses (81% of cells). This study provides a suggestion of a processing hierarchy in STS with more frequent gaze responses at anterior sites. However, the gaze responses were considerably smaller than those reported in previous work (Perrett et al., 1985), appearing mostly as deflections off a large response to a preferred head view.

Finally, there is evidence that macaque STS plays a causal role in gaze discrimination performance. Macaques with STS ablations are impaired at discriminating gaze direction across head views (Campbell, Heywood, Cowey, Regard & Landis, 1990; Heywood & Cowey, 1992). The specificity of such deficits has been called into question by the finding that ablations of the length of STS causes more general deficits in view discrimination also of non-face stimuli (Eacott, Heywood, Gross & Cowey, 1993). However, general deficits after extensive STS ablations are unsurprising since face cells represent only a subset of visually-responsive cells in STS. More specific STS ablations may become possible if ablation methods are targeted to specific subregions defined with macaque fMRI (Tsao, Freiwald, Knusten, Mandeville & Tootell, 2003).

1.4.3 Topography of STS head view cells

Most studies of head view and gaze responses in STS have sampled cells in middle and anterior STS (see Figure 1.3 for data on face-selective cells). However, Harries & Perrett (1991) sampled cells more extensively along the length of STS and reported that head view and body cells could be found in multiple clusters of approximately 3 mm size in the anterior and posterior upper banks of STS. More recent studies have used fMRI to guide single unit recording to face-selective patches in STS. Head view cells with unimodal view tunings were reported in middle lateral face patch (ML)/middle fundus face patch (MF) (Tsao et al., 2006) and cells with bimodal tunings in anterior lateral face patch (AL) (Freiwald & Tsao, 2010). Although macaque fMRI studies have found face-selective patches also in posterior STS (Pinsk, DeSimone, Moore, Gross & Kastner, 2005; Pinsk, Arcaro, Weiner, Kalkus, Inati, Gross & Kastner, 2009; Tsao et al., 2003), there have been few attempts to study head view or gaze cells in posterior STS (with the exception of Harries & Perrett, 1991). It also remains unclear whether head view and gaze cells coincide with face-selective areas defined in fMRI. However,

the size of the fMRI face patches appear broadly comparable in size to the head view cell clusters reported by Harries & Perrett (1991). Consistent with this, one macaque fMRI study found that a middle STS area was more activated when the monkeys were engaged in following turned heads to targets as compared to when they used the heads' iris colour to determine which target to saccade to (Kamphuis, Dicke & Thier, 2009). This head following effect appeared roughly in the location of the face-selective MF, although the authors did not include face localiser scans.

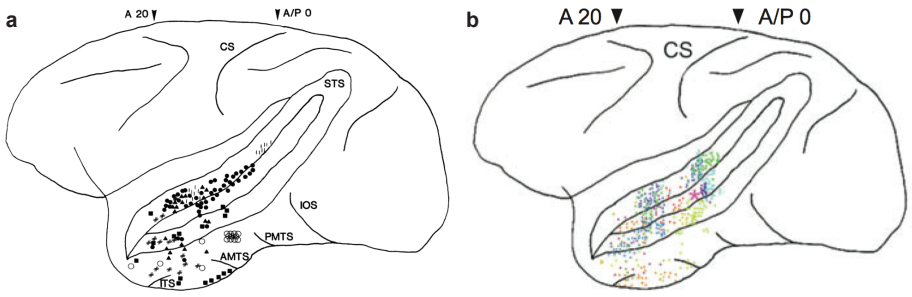


Figure 1.3: Face cells in the macaque temporal lobe as illustrated in two meta analyses. (**a**) Adapted from Perrett et al. (1992). (**b**) Adapted from Tsao et al. (2008). The location of the fMRI-defined MF is highlighted by the pink asterisk. Original studies are coded in symbols (left) or colours (right). The plotted studies in the two analyses are partially overlapping.

Beyond organisation of STS cells sensitive to faces relative to other objects on the order of millimetres, there is also evidence for a topography for specific head views on smaller spatial scales. Cells with similar head view preferences were more likely to be found at similar locations along the recording track than expected by chance (Perrett et al., 1984). Similar topographic organisations were also reported for cells selective for dynamic head turns (Perrett et al., 1985) although a later study failed to find such clustering (Hasselmo et al., 1989). The size of these 'clumps' ranged from 0.5-2 mm. However, this estimate is likely to be coarse since the limited sampling in electrophysiology makes it challenging to characterise topographic organisations. More comprehensive sampling at fine spatial scales can be obtained using optical imaging, which measures colour shifts of the cortical surface to indicate increases in cerebral blood volume caused by neuronal activity. Using this method, Wang et al. (1998) reported separate spots in anterior STS that preferred different head views. Intriguingly,

these spots were themselves organised into orderly progressions on the cortical surface from left to right profile views (Figure 1.4). Thus, there is initial evidence for patches with head view responses and organisation within the patches into separate spots that prefer different views.

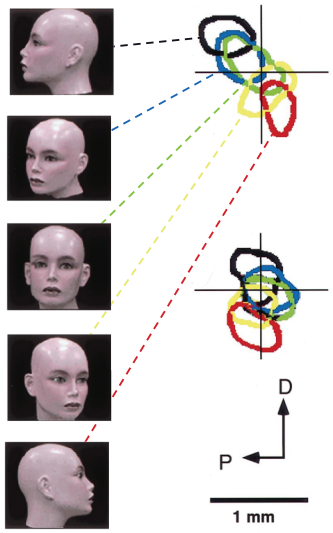


Figure 1.4: Spots with preferential responses to particular head views in macaque STS. Note that in both example cases the view preferences of the spots form an orderly topographic progression. Adapted from Wang et al. (1998).

1.4.4 Head view and gaze cells outside STS

The vast majority of electrophysiological head view and gaze studies have restricted their recordings to the STS region. However, there are a few indications of responses elsewhere in the macaque brain.

Single cells in amygdala may also be responsive to gaze direction. Brothers & Ring (1993) reported two amygdala cells that responded when the macaque viewed a video of a conspecific that made eye contact but did not extensively test whether this response was specific to eye contact. Tazumi, Hori, Maior, Ono & Nishijo (2010) tested amygdala responses to gaze more thoroughly. Such cells were tuned to direct rather than averted gaze and in some cases the gaze tuning was maintained across different head views. However, the macaques in this study were carrying out a delayed

match to sample task based on gaze direction. After 11 months of training on this task it is difficult to ascribe any neural gaze sensitivity to the percept rather than to task set confounds. Another study used fMRI to obtain evidence that amygdala responds preferentially to averted over direct gaze (Hoffman, Gothard, Schmid & Logothetis, 2007). These studies provide initial evidence for a role of amygdala in gaze perception but the exact function of these responses remains poorly understood.

Cells in the lateral intraparietal area (LIP) are also implicated in gaze perception. Many cells in LIP respond preferentially when the macaque is planning or executing saccades to a particular part of the visual field. Shepherd, Klein, Deaner & Platt (2009) located the saccadic receptive field of single LIP cells. After this, the macaques viewed gazing faces at fixation as part of an attentional cueing task. Even though the gaze cues appeared outside the cells' receptive fields, some LIP cells showed enhanced or reduced responding when the macaque perceived gaze in the direction of the saccadic receptive field. Interestingly, area LIP is implicated in attentional cueing (Gottlieb, 2007). Given the previously discussed evidence that macaques show both attentional cueing effects and micro-saccades in the direction cued by gaze (Section 1.1.1, Deaner & Platt, 2003), the key question raised by this study is whether the joint LIP responses to perceived gaze and eye movements reflect attentional cueing to the same location in space or a mirroring response where perceived and executed actions activate similar representations (Caggiano, Fogassi, Rizzolatti, Pomper, Thier, Giese & Casile, 2011).

1.5 Human neuromaging of gaze perception

With the arrival of functional imaging methods applicable to humans including positron emission tomography (PET), fMRI and magnetoencephalography (MEG), researchers have begun to explore human neural codes for perceived gaze. The findings from such studies have provided a larger-scale view of the organisation of gaze responses across cortex. Here, I discuss these human neuroimaging findings in relation to the macaque electrophysiology studies, which the human work is in many ways still catching up with.

1.5.1 Responses to gaze in STS

The first study of human neural responses to gaze was from Puce, Allison, Bentin, Gore & McCarthy (1998), who reported that posterior STS responded more to blocks of gaze shifts than to blocks of radial motion. By contrast, the motion-sensitive middle temporal area (MT) responded to both forms of motion against baseline with no significant difference between the two forms of motion. A later study found that posterior STS responds more to videos of dynamic head turns than to scrambled videos with the same low-level content (Lee, Andrews, Johnson, Woods, Gouws, Green & Young, 2010). There is thus some evidence that posterior STS responds preferentially to motion that signals attentional shifts whether this is conveyed by gaze shifts or head motion.

These studies are somewhat unusual in that regions coding gaze or head view were identified relative to a non-face control condition. Such comparisons do not isolate responses to the perceived attentional shift from general category-specific responses to faces relative to other stimulus categories. This is unfortunate because a similar posterior STS area also responds more to static images of faces than to images drawn from other categories (Kanwisher, McDermott & Chun, 1997; Kanwisher, 2010). Hoffman & Haxby (2000) avoided this problem by comparing responses to the same face stimuli in two contexts, where participants carried out a one-back detection task based on either gaze direction or identity. Regions that responded more to faces than to scrambled pictures were identified in posterior STS and lateral fusiform gyrus. Subsequent comparisons showed that posterior STS responded more during the gaze task than during the identity task, while the opposite was true in lateral fusiform gyrus. Thus, posterior STS appears to be involved more in processing gaze than in processing identity.

1.5.2 Responses to gaze relative to other face comparisons

Responses to gaze can also be isolated by comparing direct and averted gaze directly or by comparing gaze shifts to another face manipulation such as the eyes closing. Nummenmaa & Calder (2009) reported a meta-analysis of previous studies that made such comparisons (Figure 1.5).

This meta-analysis emphasises the importance of posterior and anterior STS in gaze processing. However, the inclusion of any study that reported an effect of gaze regardless of its direction obscures a considerable degree of inconsistency in the lit-

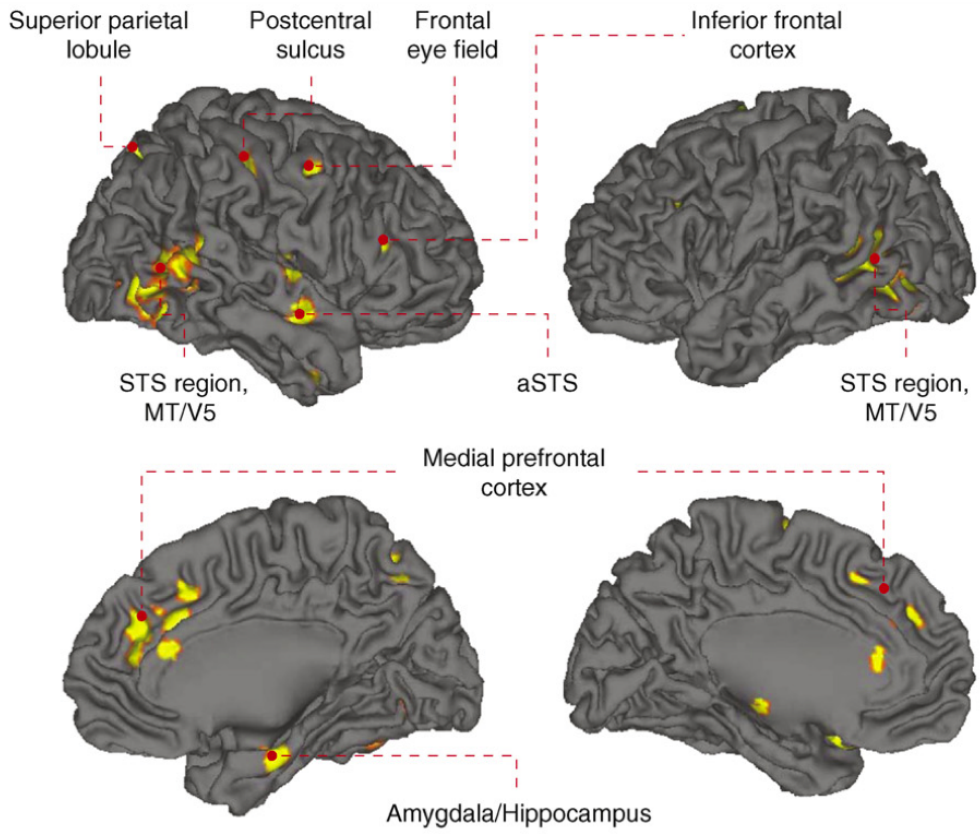


Figure 1.5: Activation likelihood estimation meta-analysis of 16 fMRI and PET studies that reported comparisons between direct and averted gaze, or between gaze shifts and another face change, such as the eye closing. Note that some studies contribute data from multiple contrasts and no distinction is made between preferential responses for direct or averted gaze directions. Effects are shown corrected for multiple comparisons using a FDR of 5%. Adapted from Nummenmaa & Calder (2009).

erature. For instance, some studies find that posterior STS responds preferentially to direct over averted gaze (Pageler, Menon, Merin, Eliez, Brown & Reiss, 2003; Pelphrey, Viola & Mccarthy, 2004; Calder, Lawrence, Keane, Scott, Owen, Christoffels & Young, 2002), while others find the opposite pattern (Engell & Haxby, 2007; Hoffman & Haxby, 2000) or a general response to gaze that does not distinguish direct from averted directions (Wicker, Henaff & Decety, 1998). Effects in anterior STS are less consistently reported.

These inconsistent effects suggest that the relative preference for direct or averted gaze in posterior STS may depend on contextual factors in the particular experiment (I return to this point in the following section). The comparison between direct and averted gaze is in itself a problematic proxy for representations of gaze direction. As discussed previously, direct gaze has particular salience (Section 1.1.2, Senju & Johnson, 2009) so any difference between direct and averted gaze may simply reflect differential engagement of mechanisms involved with for instance the attentional capture effects of eye contact (Emery, 2000; Shepherd, 2010), rather than a representation of the direction of gaze as such.

Comparisons between different averted gaze directions might more convincingly identify direction-specific representations of gaze direction since such comparisons avoid the confounding influence of eye contact effects. Studies that included multiple averted gaze directions in conventional designs either do not report comparisons (Pelphrey, Singerman, Allison & McCarthy, 2003), or find no significant difference between the averted directions (Lee et al., 2010). The only previous report of a distinction between different averted gaze directions was obtained using fMRI adaptation, a method that can detect selective tunings in intermixed neuronal populations (Andrews, 2005). Calder, Beaver, Winston, Dolan, Jenkins, Eger & Henson (2007) adapted participants to left or right averted gaze and analysed fMRI responses to probe stimuli that were consistent or inconsistent with the adapted direction. Regions in right anterior STS and inferior parietal lobule showed direction-specific adaptation to gaze. For instance, adapting to left gaze produced a reduced response to left but not to right probes. The site of this effect in anterior STS is inconsistent with the previous studies that reported differences between direct and averted gaze in posterior STS but is consistent with the typical recording site in the macaque electrophysiology studies of STS (Section 1.4.3, Perrett et al., 1992).

1.5.3 Contextual effects on STS gaze responses

The previous section described how the strongest evidence for direction-specific coding of gaze is from anterior rather than posterior STS. What is the functional role of gaze responses in posterior STS? One possibility is that responses to gaze in posterior STS reflect contextual factors. Right posterior STS responds more when an actor gazes away from an obvious target compared to when gaze is directed toward it (Pelphrey et al., 2003). Similarly, right posterior STS responds more when an actor reaches to grasp a target that he just looked away from compared to when gaze was directed towards it (Pelphrey et al., 2004). Indeed, the latter incongruence effect between gaze direction and reaching is also modulated by emotion: if an actor looks at a target, expresses disgust and then reaches for the target, posterior STS responds more than when the actor reaches for the target that was not negatively evaluated (Vander Wyk, Hudac, Carter, Sobel & Pelphrey, 2009). These studies demonstrate that responses in posterior STS are poorly predicted by the direction of perceived gaze. Instead, posterior STS appears to be coding expectancy violations based on the actor's perceived intentions (Pelphrey & Morris, 2006; Pelphrey & Carter, 2008).

A similar posterior STS region also responds more when a walker disappears for an unnaturally long time behind a book case compared to when he does not (Saxe, Xiao, Kovacs, Perrett & Kanwisher, 2004). This represents a clear expectancy violation without any manipulation to gaze direction. The cortical regions surrounding posterior STS and the adjacent temporoparietal junction are functionally heterogeneous with proposed roles in a range of functions including attentional reorienting, experiencing empathy, theory of mind, personal agency, audiovisual integration and biological motion perception (for reviews, see Decety & Lamm, 2007; Hein & Knight, 2008). Initial attempts have been made to dissociate some of these functions. For instance, theory of mind and attentional reorienting may produce distinct activations (Scholz, Triantafyllou, Whitfield-Gabrieli, Brown & Saxe, 2009), as do theory of mind and biological motion perception (Gobbini, Koralek, Bryan, Montgomery & Haxby, 2007). It remains possible that gaze responses in posterior STS also reflect distinct sub-regions from those mediating other responses in this region but no within-subjects tests of this possibility have been reported.

1.5.4 Responses to head view and gaze in STS and other regions

Much like its responses to direct versus averted gaze, posterior STS also distinguishes front-facing and turned heads (Lee et al., 2010; Natu, Jiang, Narvekar, Keshvari, Blanz & O'Toole, 2010). Since the eyes were visible in these experiments and gaze was congruent with head view, one might suppose that such responses reflect a representation of gaze direction rather than head view as such (cf. Section 1.4.2). However, published accounts that looked for gaze responses across head view changes do not support this hypothesis. George, Driver & Dolan (2001) found no effects of gaze direction or head view in STS, but reported that the fusiform gyrus responded more to direct than to averted gaze, whether the comparison was between averted and direct gaze in a front-facing or in a turned head. However, these effects did not replicate fully in a follow-up study with a larger sample (18 rather than 7 participants). Pageler et al. (2003) found no effects of gaze direction in a similar whole-brain analysis as that used by George et al. (2001) and instead observed STS effects for averted over front-facing heads. As an aside, note that Lee et al. (2010) instead found that posterior STS responded more to front-facing than turned heads, suggesting that the relative preference for front-facing or turned heads in posterior STS may be as inconsistent as its preference for direct or averted gaze (cf. Section 1.5.2). Pageler et al. (2003) also defined face-selective functional regions of interest in posterior STS and fusiform face area (FFA) and observed that FFA responded more to direct than to averted gaze in a front-facing head but showed no gaze effect in turned heads.

Contrary to the initial evidence for head view-invariant responses to gaze in macaque STS (Section 1.4.2, Perrett et al., 1985), there is only inconsistent evidence for such view-invariance in human imaging studies. When reported, view-invariant effects appear in the fusiform gyrus rather than in STS. This suggests that human STS codes gaze differently from macaque STS. Alternatively, current fMRI methods may have failed to capture neural differences between social cues that are more apparent in electrophysiology experiments.

1.5.5 Gaze responses outside STS

Like the macaque electrophysiology studies, most human imaging work has investigated gaze responses with particular focus on STS. However, gaze effects also appear

in a range of other regions including parietal cortex, amygdala and multiple prefrontal regions (Figure 1.5). Some of the effects of gaze in parietal and prefrontal regions may reflect attentional cueing and re-orienting responses brought about by perceived gaze. One meta-analysis that compared effects of perceived gaze to effects of saccadic eye movements and attentional cueing found overlapping effects for these tasks in right intraparietal sulcus and lateral precentral gyrus (Grosbras, Laird & Paus, 2005). Such representational overlap between perceived gaze and eye movements in parietal regions matches the overlap between these processes found in single cells in macaque LIP (Section 1.4.4, Shepherd et al., 2009).

Recent human and macaque fMRI studies have also reported face-selective areas in prefrontal cortex (Chan & Downing, 2011; Tsao & Livingstone, 2008). There are initial indications that the reported face-selective area in human inferior frontal gyrus may play a role in gaze perception. Face-selectivity in this area depended on the presence of the eyes: the area responded similarly to faces as to isolated eye regions and no preferential responses to faces relative to control stimuli was observed for faces with obscured eye regions (Chan & Downing, 2011). Although these findings suggest a particular role for the eyes in guiding face responses in this region, direct comparisons between responses to different gaze directions have not yet been reported.

1.6 Outstanding questions

This chapter has described a range of approaches to the study of gaze perception, including behavioural discrimination and adaptation, macaque electrophysiology and human brain imaging. In this section I emphasise some of the recurring themes that appear across these different lines of inquiry.

The vast majority of the discussed studies used static images as stimuli, even though gaze shifts and head turns are clearly dynamic in nature outside the laboratory. There is initial evidence that neural responses to dynamically turning heads in macaque anterior STS are distinct from responses to static stimuli. In the human literature, there have only been isolated attempts to study responses to such dynamic cues with initial evidence that posterior rather than anterior STS may be implicated (Lee et al., 2010).

Perhaps the most striking feat of human gaze discrimination is our ability to per-

ceive that visually distinct stimuli convey the same direction of attention. Nevertheless, relatively little is known about how this is implemented in the brain with most macaque electrophysiology studies manipulating head view only and most human imaging studies manipulating gaze direction in the context of a front-facing head only. A lingering concern in such studies is that apparent sensitivity to the social cue (gaze direction, head view) may reflect simple coding of the retinotopic or featural information that conveys this information whether this is the outline of a head profile or a change to the iris eccentricity in a gaze shift. By identifying representations that code gaze direction across such low level visual differences a more convincing case may be made that the responses really reflect a representation of gaze direction as such.

More generally, the human imaging work has produced conflicting results. For example, it is difficult use the literature to make firm predictions about which regions code the direction of gaze or head view. Most of these studies used fMRI data that had been acquired at a resolution of, at best, 3 mm isotropic voxels, spatially smoothed by a Gaussian kernel of, at best, 6 mm and often also subjected to spatial normalisation to a template space. Given that the macaque results suggest that STS representations of head view are organised on the order of millimetres, it is perhaps unsurprising that measures obtained at such resolutions have produced mixed effects. If imaging studies of human gaze perception are to catch up with the evidence from macaque electrophysiology it may be necessary to use methods that are more sensitive to fine-grained organisation, including fMRI adaptation and MVPA.

A related concern is that human-macaque comparisons are often confounded with differences between fMRI (the typical human method) and electrophysiology (the typical macaque method). Human fMRI studies are often justified on the basis of previous evidence from macaque electrophysiology so it is essential to validate such inferential leaps by showing that applying similar methods to both species produces similar results. The arrival of macaque fMRI provides a promising opportunity to make species comparisons without also introducing fundamental method differences.

1.7 Organisation of the current thesis

This thesis applies MVPA to fMRI data (a method described in Chapter 2) in order to better understand neural representations of perceived gaze direction. Experiments

1-4 (Chapter 3) present the first application of MVPA to studying direction-sensitive representations of dynamic head turns in the human brain. I observed a region in anterior STS/superior temporal gyrus (STG) that distinguished head turn direction but not direction of motion in rotating ellipsoid control stimuli. Experiments 5-6 (Chapter 4) explored the representational content of such direction-sensitivity by manipulating both gaze direction and head view in images of faces. I found that a region in anterior STS codes gaze direction in a manner that is invariant to head view, thus suggesting that human anterior STS represents the direction of another's attention regardless of the exact cue configuration that conveys this. In Chapter 5, I apply similar methods as in Experiment 1 to investigate direction-specific representations of gaze direction (Experiment 7) and head view (Experiment 8) in the macaque brain with fMRI. These experiments form a bridge between the single unit data from macaque electrophysiology that form the key empirical background for the thesis and the human fMRI data in Experiments 1,4 and 5. Chapter 6 offers concluding remarks on the neural representation of gaze direction in man and macaque and proposes directions for future research.

Chapter 2

Using MVPA of fMRI data to study gaze perception

Many of the key results in this dissertation were based on various forms of MVPA. Unlike conventional univariate fMRI analysis where each voxel is analysed separately, MVPA methods combine evidence across voxels in order to enhance sensitivity. The multivariate method by which the voxels are combined can differ considerably. Here I focus on a machine learning approach where a classifier learns to distinguish voxel patterns evoked by different conditions and a representational similarity approach that proceeds in two steps: first, the spatial correlation across voxels is computed as an index of how similar the responses to different conditions are; second, these observed similarities are compared to hypothesised similarity structures that are generated according to different predicted representations.

The chapter begins with a discussion of previous MVPA results that suggest the benefits of using MVPA to study gaze direction representations. I then describe the MVPA methods I applied in the reported experiments. In particular, I justify the various analysis choices in greater detail than is offered in the method sections for the relevant experimental chapters. In the interest of brevity I assume some familiarity with fundamental fMRI physics and conventional univariate fMRI analysis methods.

2.1 The applicability of MVPA to gaze perception

There is evidence that MVPA methods are more sensitive than conventional univariate analysis methods and more controversially, such benefits may originate in sensitivity to fine-grained information, including organisation at the scale of cortical columns. In this section I discuss this evidence with particular emphasis on how the findings relate to what is known of the cortical organisation of STS head cells.

2.1.1 Enhanced sensitivity of MVPA relative to univariate analyses

The first study that applied MVPA methods to fMRI data reported a relatively simple discrimination between different object categories in IT (Haxby, Gobbini, Furey, Ishai, Schouten & Pietrini, 2001). Similar comparisons also produce effects in univariate analyses (Kanwisher et al., 1997; Kanwisher & Yovel, 2006; Kanwisher, 2010) so the MVPA effects were used to argue in favour of distributed versus localised organisation of category-selectivity; this debate is ongoing (Spiridon & Kanwisher, 2002; Hanson, Matsuka & Haxby, 2004; Hanson & Halchenko, 2008).

Subsequent studies have used MVPA to report increasingly subtle distinctions between conditions that do not usually produce differences in univariate analyses at similar statistical significance criteria. In two studies, Kamitani & Tong (2005, 2006) demonstrated that MVPA could be used to discriminate line orientation and direction of motion in multiple early visual areas. These effects indicated that MVPA may enable neuroscientists to ask more specific questions at a subordinate or within-category level where conventional univariate fMRI methods usually fail. A study that applied such within-category comparisons to face perception found that a classifier could distinguish which of two faces had been presented based on response patterns in anterior IT but not based on patterns in FFA, even though the latter region responds more to faces than to non-faces (Kriegeskorte, Formisano, Sorger & Goebel, 2007). This was a surprising finding because researchers who focus on univariate face-selectivity had argued that areas that are selective to faces relative to non-face objects are the same areas that support discrimination of different faces (Kanwisher, 2000; Grill-Spector, Knouf & Kanwisher, 2004; Kanwisher & Yovel, 2006). The fact that this does not seem to be the case highlights the necessity of identifying areas that show within-category effects.

Such comparisons often appear to require MVPA methods to produce statistically significant effects.

There are many other demonstrations of how statistically significant MVPA effects may be obtained for comparisons that do not typically produce comparable univariate effects. For instance, response patterns in early visual cortex can be used to discriminate colour (Seymour, Clifford, Logothetis & Bartels, 2009), the dominant percept during binocular rivalry on a trial-by-trial basis (Haynes & Rees, 2005), or indeed to reconstruct perceived images (Miyawaki, Uchida, Yamashita, Sato, Morito, Tanabe, Sadato & Kamitani, 2008). Such enhanced sensitivity makes MVPA an attractive method for studying gaze direction representations where the aim is to make within-category face comparisons (cf. Kriegeskorte et al., 2007) that have produced weak or inconsistent effects in previous univariate studies (Section 1.5.2).

2.1.2 The spatial frequency of MVPA effects

The discussion above suggests that MVPA outperforms univariate analysis methods but the exact reason for this advantage is controversial. A significant MVPA effect in a region could in principle reflect a purely univariate difference, where each voxel shares the same mean effect with added noise. In this case, MVPA should perform no better than a univariate analysis of the regional average. The fact that MVPA frequently performs better than univariate analyses of the mean time course suggests that information is coded in spatial patterns within the analysed region and not only in the regional mean (in fact, many MVPA methods remove the mean before analysing patterns, as is done in correlation-based approaches). At what scales are such effects obtained? This discussion focuses primarily on representations of line orientation in primary visual cortex, which has become the test case for debates concerning the spatial scale of MVPA effects.

2.1.2.1 Hyper-acuity

Early attempts to explain MVPA effects argued that the effects originate in hyper-acuity, that is, sensitivity to information that is coded at a scale beyond the conventional resolution limit of fMRI (Kamitani & Tong, 2005; Haynes & Rees, 2006). For instance, ocular dominance columns in human primary visual cortex have a width of

approximately 0.5 mm on the cortical surface (Gardner, 2010), with is beyond the Nyquist frequency for the typical fMRI acquisition (3 mm isotropic voxels). Pattern effects can still be driven by such sub-voxel signals if the distribution of columnar preferences is uneven across voxels so that the voxel mean has a weak bias in favour of the columnar preference that happens to be slightly more frequent in that voxel. In this view, the MVPA effect is a consequence of MVPA methods pooling the evidence from such weak preferences into a strong prediction (Figure 2.1).

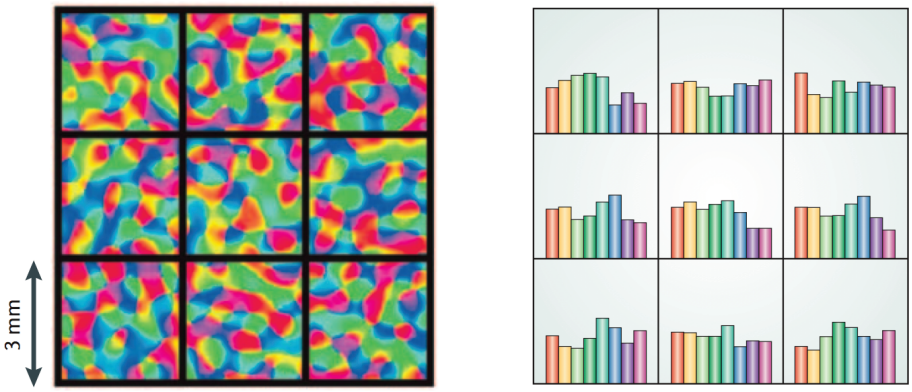


Figure 2.1: Demonstration of the hyper-acuity account of pattern effects. The left panel shows a schematic illustration of columnar orientation preferences in primary visual cortex. The right panel shows how random biases in the distribution of columnar preferences results in subtle preferences for certain orientations if each voxel's response level is determined by these distributional biases. Colours represent different orientation columns. Adapted from Haynes & Rees (2006).

The hyper-acuity account has subsequently been criticised in two ways. First, Op de Beeck (2010) reported that 8 mm spatial smoothing did not disrupt classification of line orientation in early visual cortex, or of object category in IT. This may seem inconsistent with the hyper-acuity account since smoothing means that each voxel now samples putative subtle biases across a larger area of cortex, thus producing weaker deviations from the mean. However, this argument fails to consider that smoothing also reduces high spatial frequency noise (Kamitani & Sawahata, 2010). The hyper-acuity account makes no prediction for how smoothing or voxel size should affect MVPA effects since both univariate and MVPA effects also depend on signal to noise ratio (Bodurka, Ye, Petridou, Murphy & Bandettini, 2007). Thus, this critique does not

necessarily pose a problem for the hyper-acuity account.

However, MVPA effects also remain in the presence of moderate participant movement (Kriegeskorte, Bandettini & Cusack, 2009) or even across scanner sessions (Kamitani & Tong, 2005). This is more problematic for the hyper-acuity account because any movement that approaches half a voxel should be catastrophic if each voxel is sampling an idiosyncratic bias of columnar tunings: consider the consequences of a transformation to the voxel grid in Figure 2.1. Subsequent theorists have accommodated these findings by considering how fMRI voxels sample neural activity via the cortical vasculature.

2.1.2.2 Spatiotemporal filtering through the neurovasculature

Because standard blood oxygen-level dependent (BOLD) fMRI measures neural activity indirectly through the vasculature, the measure at a given voxel is unlikely to be completely independent of its neighbours. Even focal neural activity will produce BOLD responses millimetres away from the origin (the BOLD point-spread function exceeds 3 mm, Op de Beeck, 2010) and BOLD responses measured at large draining veins may reflect even more distal neural events (Shmuel, Chaimow, Raddatz, Ugurbil & Yacoub, 2010; Gardner, 2010). The key argument from spatiotemporal filter theorists (Gardner, 2010; Kriegeskorte et al., 2009) is that such apparent distortion may be the very reason why MVPA can detect effects from columnar organisations. The vasculature may act as a complex filter that aliases high spatial frequency columnar organisation into low spatial frequency vascular signals. Such vascular organisation may be on a scale that exceeds voxel size by a wide enough margin to survive realignment errors from participant motion or the combination of data from different scanner sessions (Kriegeskorte et al., 2009).

A key claim in this account is that there is some systematic organisation between neuronal preferences and vascular organisation. This organisation may be a consequence of random variation in how veins and capillaries sample columns (Kriegeskorte et al., 2009). Alternatively, the organisation may arise because columns with shared tunings often co-activate and the resulting correlated metabolic requirements encourage the development of shared vascular inputs. This process would effectively magnify hemodynamic differences between features coded by different columns (Gardner,

2010). Such proposals are speculative at present but they present a plausible account for why MVPA of coarse-scale fMRI data can be sensitive to effects that are likely driven by sub-voxel columnar organisation.

2.1.2.3 Estimating pattern scales

An alternative to hyper-acuity and spatiotemporal filter accounts is to explain line orientation MVPA effects through large-scale biases in favour of certain orientations. Primary visual cortex responds more to cardinal than to oblique orientations (the oblique effect, Furmanski & Engel, 2000) and more to orientations where the lines pass through the fovea compared to orientations where the lines do not (the radial bias, Sasaki, Rajimehr, Kim, Ekstrom, Vanduffel & Tootell, 2006). If such large-scale biases explain line orientation MVPA effects there is no need to invoke complex theories for how fMRI achieves sensitivity to sub-voxel topographic organisation.

Mannion, McDonald & Clifford (2009) showed that MVPA can be used to discriminate the orientation structure in spiral glass patterns. Such patterns are made up of many small segments and have no large-scale organisation. It is thus unlikely that classification of such patterns could originate in any large-scale retinotopic biases such as those previously discussed. In a subsequent study, Swisher, Gatenby, Gore, Wolfe, Moon, Kim & Tong (2010) estimated the spatial scales at which line orientation MVPA effects are coded. The authors applied high-pass and low-pass filtering to high resolution fMRI responses to line orientation in human and cat primary visual cortex. The cat data was acquired at a sufficient resolution to image individual orientation columns and orientation discrimination in this dataset remained above chance after high-pass filtering at the scale of individual columns (0.5 mm). This indicates that columnar-scale organisation made a contribution to the MVPA effect. Effects in both species also survived low-pass filtering of up to 10 mm suggesting that orientation information was present at a range of spatial scales. Patterns at scales beyond 10 mm made a relatively weak contribution to the effects so large-scale biases did not contribute to the effects.

More recently, however, Freeman, Brouwer, Heeger & Merriam (2011) provided evidence that this type of spatial filtering fails to isolate fine-grained or columnar signals. High- and low-pass filtering produced near-identical effects on classification of

line orientation as on classification of retinal location, even though retinal location is organised at a supra-voxel scale in visual field maps. Furthermore, the authors used phase mapping of line orientation to reveal a large-scale topographic map of line orientation, which was in register with the visual field map. Thus, line orientation has proved to be an unfortunate test case for investigations of the spatial scale of MVPA effects because there seems to be a large-scale organisation of this particular feature.

Note that these findings do not imply that MVPA cannot be sensitive to sub-voxel organisation, only that the presence of a MVPA effect cannot be taken as evidence for the existence of sub-voxel organisation (for an example of this misconception, see Hassabis, Chu, Rees, Weiskopf, Molyneux & Maguire, 2009). It is possible that other MVPA effects that were previously thought to reflect fine-scale organisation will also be shown to reflect supra-voxel topography. This suggests a fruitful approach to research, where sensitive but spatially-imprecise MVPA methods are used to reveal the presence of a new effect (e.g., discrimination of line orientation) and this finding guides subsequent univariate attempts to identify the exact topography that supports the MVPA effect.

2.1.2.4 Implications for gaze direction codes

The use of MVPA in the study of codes for gaze direction or head view is justified by two arguments. First, previous univariate fMRI studies of gaze have produced inconsistent effects, which suggests the need for more sensitive methods. There is evidence that MVPA methods achieve greater sensitivity than univariate methods. Second, responses to head view have been proposed to be organised at a columnar scale (Wang et al., 1998, Section 1.4.3) and although it remains unclear how sensitive MVPA is to such small-scale organisation, theoretical accounts suggest that such organisations may be more easily detectable using MVPA.

2.1.3 Comparing MVPA to fMRI adaptation

An alternative method for achieving enhanced sensitivity, particularly to fine-grained spatial organisations, is to use conventional univariate analyses applied to fMRI adaptation designs. Much like behavioural after-effects (Section 1.3), fMRI adaptation is based on the premise that repetition of the same stimulus should adapt neuronal pop-

ulations thus producing a reduced BOLD response in areas that represent the adapted stimulus dimension (Grill-Spector & Malach, 2001; Grill-Spector, Henson & Martin, 2006). Such methods have previously been used to obtain direction-specific adaptation to gaze direction in human anterior STS (Calder et al., 2007).

Sapountzis, Schluppeck, Bowtell & Peirce (2010) compared orientation-specific adaptation and MVPA effects across early visual areas. The magnitude of adaptation and MVPA effects were closely correlated across areas but the p values for the MVPA effects were consistently smaller, suggesting that similar information was captured by both methods although the MVPA method was more sensitive. However, this may not represent conclusive evidence for the advantage of MVPA over fMRI adaptation because Sapountzis et al. (2010) used an event-related fMRI adaptation design but a block MVPA design. This decision was justified by claiming that these designs are optimal for the respective methods. However, given that block designs have been successfully used with fMRI adaptation (e.g., Ewbank & Andrews, 2008) and rapid event-related designs with MVPA (e.g., Kriegeskorte et al., 2007), it is not clear from this study alone that the advantage of MVPA was a result of analysis differences rather than the increased design efficiency of block designs relative to event-related designs (Henson, 2003).

In any case, fMRI adaptation designs are critically limited in that adaptation effects necessitate a certain order of stimulus presentation, which limits the number of comparisons that can reasonably be made within a scan session. Multivariate pattern analysis places fewer restrictions at the experimental design stage, which makes condition-rich designs possible (I return to such designs in Section 2.2.2).

Ultimately, fMRI adaptation and MVPA methods depend on different assumptions concerning, respectively, the effects of repetition on neural representation and the spatial organisation of hemodynamic responses. The strongest case for a region's involvement in a given function is obtained if these methods provide convergent evidence despite their differing assumptions.

2.2 Multivariate pattern analysis methods in the current dissertation

In this section I describe the general features of the MVPA approach used in this dissertation. Although this section will necessarily anticipate some experimental details, the reader is referred to the following chapters for a full account of the experimental design for particular experiments.

2.2.1 Identifying MVPA effects using classifiers

The first study that applied MVPA to fMRI data used a type of classifier (Haxby et al., 2001) and the vast majority of MVPA studies that followed it used similar methods. Classifiers originate in machine learning, a field of computer science focused on algorithms that enables software to behave adaptively and learn from the environment. Classifiers carry out a wide range of tasks outside of neuroscience, including predicting consumer preferences for shopping websites, identifying images of faces for search engines and converting speech to text (Mitchell, 1997).

The fundamental concept in classifier-based MVPA is prediction. A classification algorithm learns to distinguish brain responses evoked by different experimental conditions based on a training dataset and the accuracy of the learned mapping between brain responses and experimental conditions is tested with an independent test dataset. Statistical inference is based on how well the classifier performs at predicting the correct experimental condition based on its response pattern in the test dataset. In this section, I discuss the specifics of how the data is divided into individual examples for training and testing the classifier, how the division between train and test datasets is carried out and how the classifier's performance on the test dataset is assessed.

2.2.1.1 Generating training examples for classification

The first step in classifier-based MVPA is to generate training examples based on the echoplanar imaging (EPI) volumes from the fMRI scanner. Published reports exist of many solutions, ranging from providing the classifier with a set of EPI volumes within a time window after each experimental event (e.g., Hanson & Halchenko, 2008), to fitted parameter estimates from a univariate generalised linear model (GLM) where each

example may comprise a single trial, all repetitions of a particular condition within a run (Figure 2.2B), or repetitions across multiple runs (e.g., Haxby et al., 2001). There is a fundamental trade-off here between offering the classifier many noisy training examples, or fewer but more consistent training examples (Pereira, Mitchell & Botvinick, 2009; Golland, Liang, Mukherjee & Panchenko, 2005).

Experiments 7-8 in this dissertation (Chapter 5) used a more recently developed method for extracting single trial or block estimates (Mumford, Turner, Ashby & Poldrack, 2011). In this approach, a separate GLM is fitted for each event in the experiment with one regressor for the event and a single second regressor that includes all other events in the run regardless of condition. Parameter estimates obtained from this approach produce enhanced classification relative to estimates from a GLM where all events are modelled simultaneously using one regressor per event. The authors suggested that this is because regressors for temporally-adjacent events tend to be highly correlated in GLMs with one regressor per event, which makes for variable estimates (Mumford et al., 2011).

Training and test examples must also be divided into independent sets. Most pattern classifiers always perform with above-chance accuracy if they are tested with the same data that they were trained with. In fMRI data, temporal autocorrelation of the BOLD response means that it is not sufficient to merely select different sets of examples for training and testing the classifier: the examples in different sets must also be so far apart in time that temporal autocorrelation is not a concern (Kriegeskorte et al., 2009).

Studies that use a GLM to generate training examples generally use the raw parameter estimates but better performance may be obtained if T maps are used instead (George, Johnson, Lu, Jin, Strasburger, Laken & Kozel, 2009; Misaki, Kim, Bandettini & Kriegeskorte, 2010). This is because the T statistic scales the parameter estimate by its variance, thus downscaling the contribution of noisy voxels. If T maps are used it is essential that separate GLMs are fitted to each set because independence between sets will be violated if the variance estimate spans all sets. However, if parameter estimates are used as training examples it is sufficient to use estimates from a single GLM as long as the modelled events for each regressor are sufficiently separated in time to yield independent estimates.

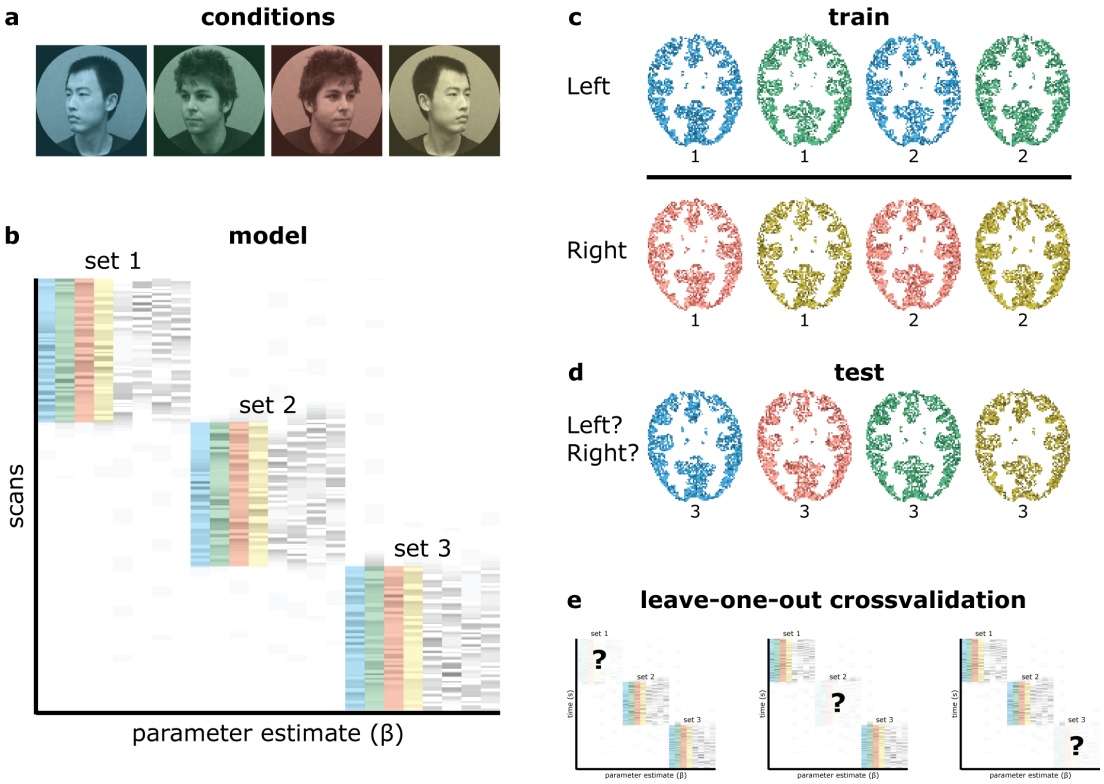


Figure 2.2: Estimating crossvalidated classification accuracy (based on a simplified account of the method in Experiment 1, Chapter 3). (**a**) The experiment included four relevant conditions: two faces turning their heads leftward and rightward. (**b**) These stimuli were presented during a single run in a rapid event-related design. Separate regressors were fitted for each stimulus and for each set of the experiment. This produced two training examples per head turn direction and set. (**c-d**) Example of leave-one-out crossvalidation where the first two sets form the training data and the third set forms the test data. (**e**) This procedure is repeated for each possible combination of training and test data; the average performance across all combinations is the final accuracy estimate.

2.2.1.2 Defining independent sets of examples

The independent set in fMRI experiments is generally a scanner run, or a subset of trials from a single run with sufficient temporal separation between preceding and following subsets. The optimal split between training and test data is leave-one-out cross-validation, where the classifier is trained on the n unique combinations of the $n - 1$ independent sets and accuracy is tested on the examples in the remaining set. The final accuracy is the average of the performance for the n train/test combinations (Figure 2.2). The benefit of this method is that each test of the classifier is based on all available independent training data, as opposed to alternative splits such as the odd-even split (Pereira et al., 2009).

One concern with set-based crossvalidation is that the classifier has no knowledge of the order in which the examples occurred in the experiment and is thus unable to take account of order effects during the scanner session. Such effects may result from, for instance, the participant's head translating slowly during the scanner session or long-term adaptation to the experimental stimuli. Some studies minimise order effects by generating one set of examples as the average for the data from even sets and another as the average from the odd sets (e.g., Haxby et al., 2001; Kriegeskorte et al., 2007). Any order effects between the two sets are thus greatly reduced. I have implemented an alternative solution where the examples in each set from the first half of the experiment are averaged with the examples from the corresponding set in the second half of the experiment. For example, in an experiment with 6 sets, averaging produces a dataset with 3 sets. Each example in set 1 of the averaged dataset is an average of the examples across sets 1 and 4 of that condition in the full dataset, each condition in set 2 of the averaged dataset is an average of the examples across sets 2 and 5 in the full dataset and finally, each condition in set 3 is an average of the responses across sets 3 and 6. Leave-one-out crossvalidation can then proceed as normal based on the averaged sets. This solution reduces session effects through averaging but preserves more training examples than the odd-even average method for designs with 6 or more initial sets.

Beyond order effects, there are many other ways in which inhomogeneities in the fMRI data can disrupt classification. Most classifiers are sensitive to such common occurrences as voxels with particularly large responses or variance, or global reductions in responses across all voxels from one set to the next, perhaps caused by participant

fatigue. Classification performance is thus often enhanced if the training examples in each set are separately Z-scored across voxels (Hanson & Halchenko, 2008; Chen, Pereira, Lee, Strother & Mitchell, 2006; Pereira et al., 2009) to ensure that the distribution of voxel responses in each set is centred on 0 with a standard deviation of 1. It can be helpful to precede Z-scoring by Winsorising (capping) outliers to for instance 2 standard deviations from the mean (Seymour et al., 2009). If this is not done Z-scoring could introduce inhomogeneities across sets as some sets may contain extreme outliers that shift the set mean and standard deviation while others may not.

2.2.1.3 Training and testing the classifier

All classifier analyses in this dissertation were carried out using a linear support vector machine (SVM) classifier. This classifier is among the most popular in the field and in methodological evaluations it matches or outperforms other classifiers (Mitchell, Hutchinson, Niculescu, Pereira & Wang, 2004; Pereira & Botvinick, 2011; Chen et al., 2006; Mourao-Miranda, Bokde, Born, Hampel & Stetter, 2005; Misaki et al., 2010), although performance is often quite similar across classifiers for fMRI data. To intuit how the SVM classifier is trained, imagine an n-dimensional space, where n is the number of voxels in the analysed region. Training proceeds by identifying a hyperplane that separates the training examples belonging to the two conditions by maximising the margin between the examples closest to the plane (the support vectors) and the plane itself (Figure 2.3; for conceptual introductions, see Pereira et al., 2009; Noble, 2006; Mur, Bandettini & Kriegeskorte, 2009).

Many studies that use SVM classifiers manipulate the C parameter, which determines the hardness of the margin. A small C corresponds to a soft margin where outlier examples are allowed to fall on the wrong side of the hyperplane in the training set in the interest of arriving at a simple solution, while a large C corresponds to a hard margin where the hyperplane is placed to maximise the distance between the conditions regardless of complexity (Laconte, Strother, Cherkassky, Anderson & Hu, 2005). Soft-margin SVM classifiers are generally preferable since such classifiers are less prone to over-fitting of outlier examples. C must be set before data analysis or optimised for the data in each training split (Pereira & Botvinick, 2011). In the experiments reported here I used the default setting in the multivariate pattern analysis in

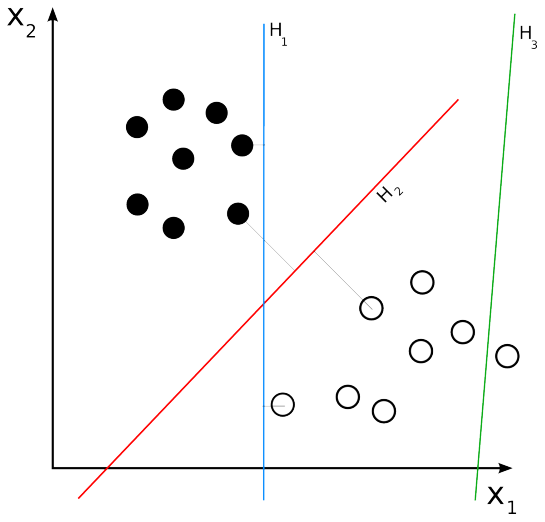


Figure 2.3: Illustration of different hyperplanes for MVPA using a linear classifier. The training examples for two conditions (black and white dots) are plotted as a function of two voxels $[x_1, x_2]$. The green hyperplane does not separate the conditions. The blue and red hyperplanes both separate the conditions but the red hyperplane achieves the greatest margin (faint grey lines) between the examples closest to the plane (the support vectors). Adapted from http://en.wikipedia.org/wiki/Support_vector_machine (retrieved 7 August, 2011).

python (PyMVPA) software implementation (Hanke, Halchenko, Sederberg, Olivetti, Fründ, Rieger, Herrmann, Haxby, Hanson & Pollmann, 2009), where C is scaled by the norm of the training data.

Once the classifier is trained, testing involves determining on which side of the hyperplane the test examples fall. During training weights are assigned to each voxel according to how well the conditions are separated along that dimension and the test stage for most linear classifiers is equivalent to the inner product of the observed data vector and the weights vector. A positive result corresponds to predicting one condition and a negative result the other condition. This predicted condition is then compared to the true condition label to assess prediction accuracy.

In most basic research applications, absolute classification accuracy conveys little meaningful information on its own. If the training examples have been averaged extensively performance may be high across a small number of tests and conversely, poor but above-chance performance across a large number of tests may indicate the presence of a consistent effect. It is therefore preferable to focus on p values rather than accuracy when assessing classification performance (Kriegeskorte & Bandettini, 2007). There are clearly many other contexts where accuracy is crucial, for instance if MVPA is used to distinguish clinical populations from healthy controls (e.g., Ecker, Marquand, Mourao-Miranda, Johnston, Daly, Brammer, Maltezos, Murphy, Robertson, Williams & Murphy, 2010) but this is very different from the null-hypothesis significance testing upon which basic research is based. p values may be obtained through binomial tests, which are exact as long as the test examples are independent (Pereira et al., 2009). Another possibility is to use permutation tests where the training examples are randomly reshuffled many times in order to build a null distribution for the dataset that actual performance is then compared to. Both methods produce similar results in methodological evaluations (Pereira & Botvinick, 2011) so in the interest of computational efficiency I use binomial tests for single-participant analyses in this dissertation.

2.2.2 Representational similarity analysis

Although classifier-based MVPA is a powerful method, its applicability to the current dissertation is limited by two factors. First, many hypotheses are not readily addressed by finding significantly above-chance classification performance between

different conditions. For instance, parametric relationships are not easily tested within a classification framework. Second, splitting the data into train/test sets is expensive and should be avoided whenever possible in order to maximise power (cf. Friston, Rotshtein, Geng, Sterzer & Henson, 2006). One solution to these limitations is to use representational similarity analysis (RSA), a correlation-based MVPA method that can accommodate parametric predictions across many conditions (Kriegeskorte et al., 2008) and does not necessarily require splitting the data. Although RSA is a very flexible analysis framework, I focus here on its application in hypothesis-testing rather than in exploratory analysis (e.g., Kriegeskorte, Mur, Ruff, Kiani, Bodurka, Esteky, Tanaka & Bandettini, 2008).

2.2.2.1 Generating representational dissimilarity matrices

The first stage of RSA is to generate a representational dissimilarity matrix (RDM), which summarises all pairwise comparisons between the conditions in the experiment according to some distance metric, such as $1-r$ (Pearson correlation coefficient) across voxels. Much like in the classifier-based studies, T maps are used as the inputs for this analysis (Section 2.2.1.1). These T maps can be based on a GLM with one regressor per condition, but if the experiment includes independent subsets it can be helpful to generate separate RDMs for each subset and then average these set RDMs to make the final RDM (Figure 2.4).

The benefit of this approach is that it removes the influence of any transformations to the response patterns across sets. Consider the case of a participant whose head translates slowly during the experiment. The pattern responses to each condition may be consistent but movement across the experiment means that the patterns have been transformed relative to the voxel grid, thus producing poor fitted responses when a single set of regressors is fitted for each voxel across all scans. Averaging RDMs across sets means that the exact voxel pattern does not have to match across sets as long as the dissimilarities between the conditions are consistent.

2.2.2.2 Testing experimental predictions

The second stage of RSA is to compare the data RDM to predictor RDMs. These can be generated according to conceptual or computational models, or by an independent

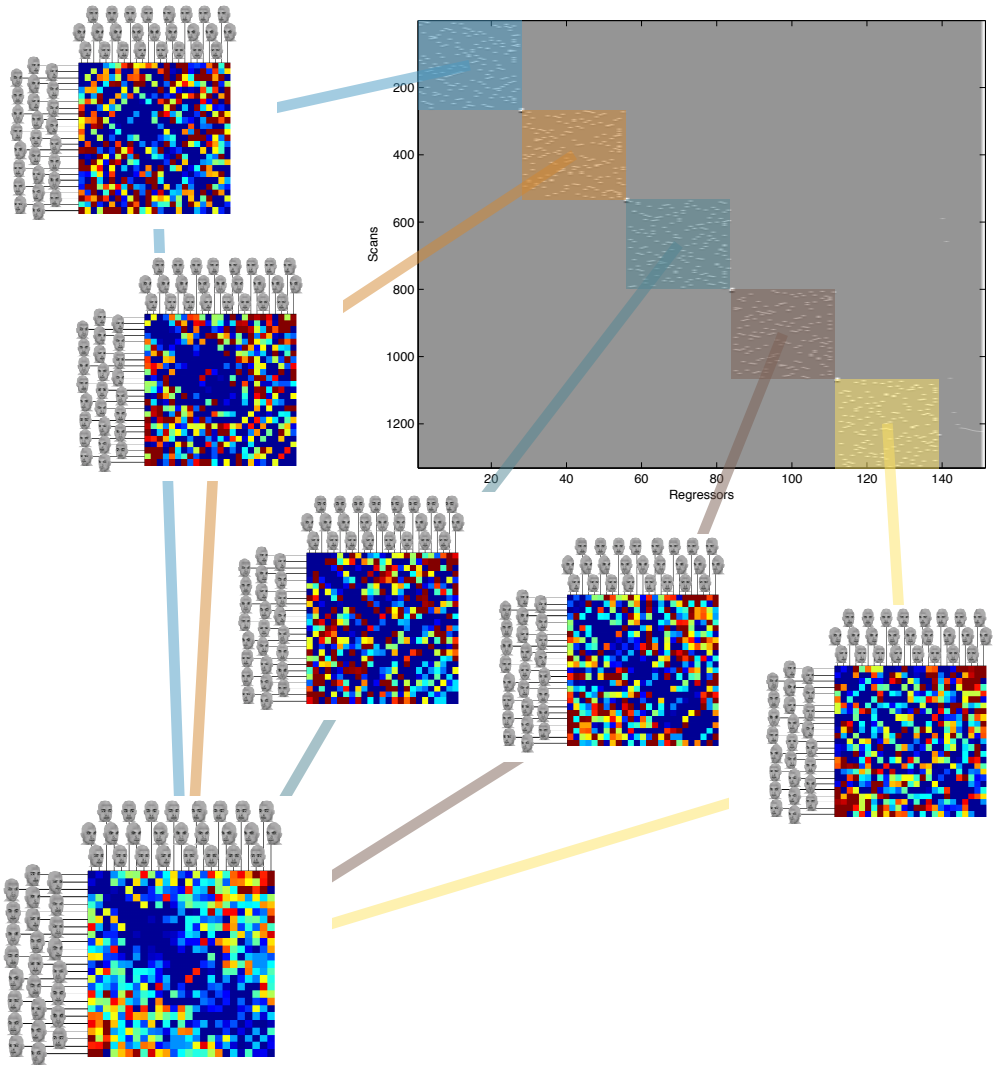


Figure 2.4: Averaging RDMs from each set of the experiment to create a final data RDM (based on a simplified account of the design in Experiment 5, Chapter 4). Separate regressors are fitted to the data for each of 25 head/gaze configurations in each of 5 independent subsets of the experiment. Subset RDMs are generated for each set ($1 - r$ across voxels in T maps contrasting each stimulus against baseline) and the unique elements of each subset RDM are averaged across subsets (lower left panel). Simulated data: cool colours correspond to small numbers (high correlations) and hot colours correspond to large numbers (low or negative correlations). Note that the use of a single GLM for all subset RDMs violates independence. However, this is not an issue if the procedure is only used to generate a mean RDM across subsets.

data source such as the participant's gaze discrimination performance (see Experiment 5, Chapter 4). The similarity between the data and predictor RDMs is measured by the Spearman rank correlation between the unique elements in each RDM. Rank correlation is preferable to Pearson's correlation because it only assumes a monotonic rather than a linear relationship between the model and data RDMs. Such monotonic relationships fit better with most experimental hypotheses. For example, in an experiment where a parametric difference between left, direct and right gaze is expected, it may not be hypothesised that the difference between left and direct should be twice the difference between left and right (the linear relationship tested by Pearson's correlation), only that the order of dissimilarities should change systematically as a function of gaze (the monotonic relationship tested by Spearman rank correlation).

If there are correlations between competing predictor RDMs it is important to show that the relationship between the data and one predictor RDM survives removing the influence of the correlation between the data and the competing predictor RDM. I decided to carry out these tests with a partial Spearman rank correlation. This is a new development in RSA. Another possibility would have been to fit the different predictors to the data RDM in a GLM but this entails linearity assumptions.

Although the result of RSA is a correlation coefficient, it would be invalid to use parametric p values corresponding to this correlation for an n equal to the number of unique elements in the RDMs. This is because the elements of the RDM are not independent observations. For instance, the correlation between condition a and condition b is dependent on the correlation between condition a and c and between b and c. Kriegeskorte et al. (2008) proposed a bootstrap resampling approach to derive confidence intervals for r by resampling the entries in the data or model RDM with replacement. However, this approach becomes very computationally expensive if it is executed over a large number of regions (see Section 2.2.3 below). A simpler solution is to restrict inferences to the group level. Even if the relationship between r and p is unknown, r (or the Fisher-transformed Z for parametric analyses) can still be used as a single participant descriptive statistic to test whether r is greater than 0 across participants (Section 2.2.3.1).

2.2.3 Localising MVPA effects using searchlight mapping

The discussion in the previous sections was based on the implicit assumption that a suitable region had already been identified for MVPA. Multivariate pattern analysis is fundamentally limited compared to univariate analysis in that the combination of evidence across voxels means that there is no basis for localising effects to particular voxels in the analysis. This could be thought of as a trade-off where sensitivity is enhanced at the cost of reduced spatial precision.

Most fMRI studies in face perception use functional localisers to define regions of interest. However, it is unclear whether macaque STS regions with cells responsive to head view or gaze direction coincide with regions that are face-selective in fMRI (Section 1.4.3) and whether human face-selectivity coincides with regions showing within-category face MVPA effects (Kriegeskorte et al., 2007). In this dissertation I did not wish to make these assumptions so in addition to such tests of functional regions of interest (ROIs) I also used searchlight mapping (Kriegeskorte, Goebel & Bandettini, 2006) to localise regions where the local voxel neighbourhood showed MVPA effects.

Searchlight mapping involves moving a spherical region through the entire imaged volume and carrying out MVPA separately in each sphere. The pattern effect (usually classification accuracy or in the case of RSA a correlation coefficient) is mapped back to the central voxel in the sphere. The end result is a searchlight map, where each voxel in the volume summarises the MVPA effect for the voxels in the local neighbourhood (Figure 2.5b). In the trade-off between sensitivity and spatial precision, searchlight mapping can be thought of as an intermediate step between univariate analysis and whole-brain MVPA: sensitivity is enhanced without sacrificing all spatial precision (the searchlight acts somewhat like smoothing with a binary rather than a Gaussian kernel).

Classifier-based searchlight mapping is more sensitive if the mapping is restricted to grey-matter voxels and subtler gains also obtain if the searchlight is defined as a 2D patch on the grey-matter surface (Oosterhof, Wiestler, Downing & Diedrichsen, 2010). Masking benefits may be even larger for correlational methods such as RSA because a classifier will likely assign a small weight to noisy white-matter voxels given sufficient training data but such weighting is not possible with the correlation coefficient.

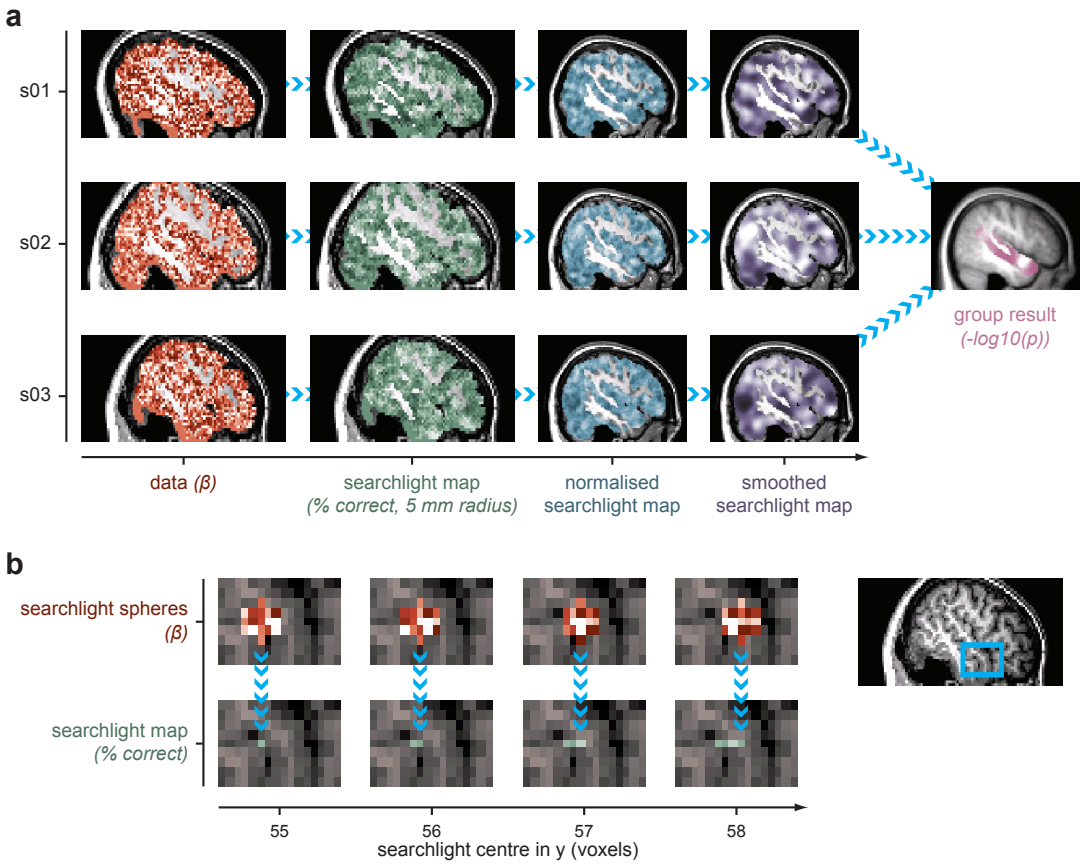


Figure 2.5: Searchlight mapping and group analysis (based on data from Experiment 1, Chapter 3). (**a**) Processing of three example participants (vertical axis). Parameter estimates in a GLM are entered into a searchlight classifier analysis. Individual searchlight maps are normalised to a common template, smoothed to correct for inter-subject alignment errors and entered into a group analysis that is restricted to the anatomically-defined right STS region. (**b**) Illustration of how a searchlight moves across the volume in overlapping steps, constructing a searchlight map of classification performance for each sphere.

2.2.3.1 Group analysis of searchlight maps

Identifying consistent group effects in searchlight maps involves normalising each map to a common template and carrying out a univariate analysis across the voxels (Kriegeskorte et al., 2007). Smoothing the normalised searchlight maps can be a helpful intermediate step because the optimal searchlight radius may be around 4 mm (Figure 2.5a, Kriegeskorte et al., 2006) and this is unlikely to produce enough low-pass filtering on its own to account for inter-subject alignment errors during normalisation and genuine anatomic variation in functional specialisation (for a previous implementation of smoothed searchlight maps, see Stokes, Thompson, Cusack & Duncan, 2009).

In this dissertation I use permutation tests as implemented in statistical non-parametric mapping (SnPM) for all group analyses (Nichols & Holmes, 2001). Permutations in SnPM are implemented by flipping the sign of the first level maps for a random subset of the participants so the only assumptions are that the first level statistic is symmetrical around 0 under the null hypothesis and that the participants are interchangeable. The latter assumption is always met in a conventional group analysis and the former assumption is reasonable for classification accuracy and correlation coefficients (accuracies are rescored to centre chance performance at 0). The test statistic in all analyses is a variance-smoothed pseudo-T statistic (10 mm full width at half maximum (FWHM)). Variance estimates are often noisier than parameter estimates in neuroimaging (Nichols & Holmes, 2001) so smoothing the variance estimate before computing the T statistic reduces the high spatial frequency components of this noise thus producing more consistent effects than with a conventional T (Figure 2.6). It would not be straight-forward to derive a parametric p for pseudo-T but this poses no problem when permutation tests are used (Nichols & Hayasaka, 2003).

Control of familywise error (FWE) is achieved in SnPM through the maximal statistic (Nichols & Holmes, 2001; Nichols & Hayasaka, 2003). At each permutation, the maximal pseudo-T in the analysed region is recorded. It can be shown that thresholding the original pseudo-T map by the pseudo-T at the $1-n$th percentile of the maximal statistic distribution controls FWE at $n\%$. If coupled with variance smoothing, this correction is more sensitive than FWE correction based on random Gaussian fields or the Bonferroni method (Nichols & Hayasaka, 2003).

Beyond this added sensitivity, there are two primary justifications for using SnPM

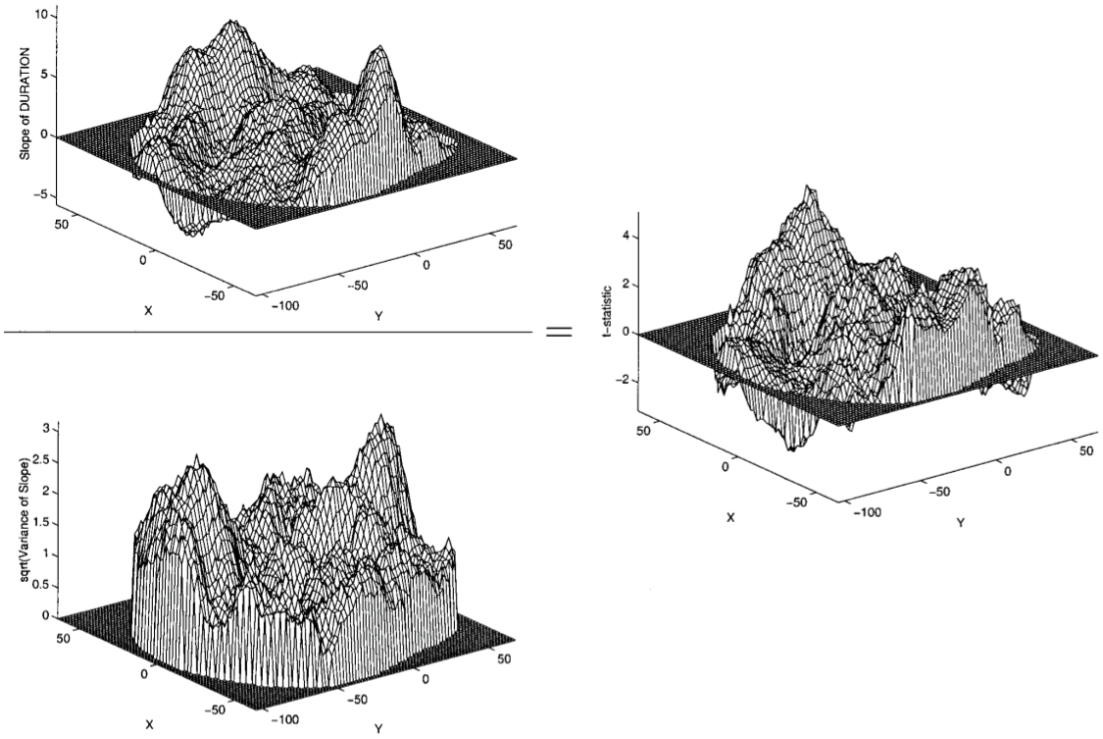


Figure 2.6: Example data from a PET experiment demonstrating the smoothness of parameter and variance estimates. Note that the parameter estimate (top left) is smoother than the variance image (bottom left) and that high spatial frequency fluctuations in the T image (right) are thus primarily caused by the variance estimate. Adapted from Nichols & Holmes (2001).

rather than conventional parametric random-effects models: first, it is desirable to avoid the assumption that the distribution of first level statistics (e.g., classification accuracy, correlation coefficients from RSA) is normal (even if Fisher-transformed). Second, because the single participant searchlight maps were grey-matter masked it is necessary to also mask the group analysis by the union of the single-participant grey-matter masks. Such discontinuous masks pose a problem for conventional FWE correction using random Gaussian fields but not for the maximal statistic approach used in SnPM.

Another sometimes reported alternative (e.g., Kriegeskorte et al., 2006; Stokes et al., 2009) is to correct group searchlight maps for multiple comparisons using the voxelwise FDR (Genovese, Lazar & Nichols, 2002). However, voxelwise FDR is too lenient in the presence of extended clusters of activation (Chumbley & Friston, 2009), which is precisely the situation for one of the experiments reported here (Experiment 1, Chapter 3).

2.2.3.2 Defining anatomical regions of interest

As the preceding chapter showed, there is considerable evidence that STS is a key region for gaze direction coding. It would therefore be excessively conservative to correct for multiple comparisons across all voxels in the grey-matter masked volume. However, it is less clear which studies should form the basis for hypothesis-derived STS regions (for instance through spherical small volume corrections centred on standard coordinates), given the mixed evidence for a role of anterior and posterior STS in human gaze codes (Section 1.5.2). I therefore decided to follow an anatomical approach in this dissertation, where the STS region is masked based in each individual study based on the sample's mean structural volume (Figure 2.7). The use of the sample mean maximises the anatomical precision of the mask at the cost of introducing an element of subjectivity in how the mask is drawn across studies. Although more objective masks may be derived from standard anatomical parcellation toolboxes such as automated anatomical labelling (Tzourio-Mazoyer, Landeau, Papathanassiou, Crivello, Etard, Delcroix, Mazoyer & Joliot, 2002) this was not possible for STS since these parcellations do not include sulci.

In all cases, analyses of the anatomical STS region were complemented with whole-

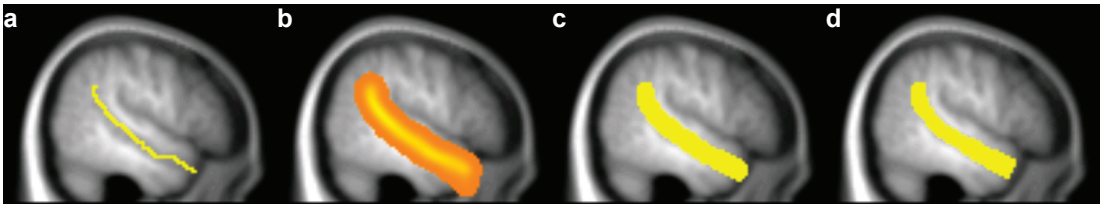


Figure 2.7: Illustration of the method by which the STS region was masked in Experiment 5 (Chapter 4). (**a**) The fundus of the STS is traced in single lines for each saggital section in the sample's mean T_1 volume. (**b**) The resulting mask is smoothed (3 mm FWHM) to expand into the upper and lower lip and to achieve a more consistent line across sections. (**c**) The smoothed mask is thresholded to restrict its extent to gyri adjacent to STS. (**d**) The thresholded mask is further restricted to regions covered by the group grey-matter mask.

brain analyses in order to capture potential effects of interest outside STS. Note that because multiple comparisons correction in SnPM is based on the maximal statistic across the analysed volume, it was necessary to carry out separate permutation tests for each of these regional analyses to generate FWE-corrected p values for each region.

Chapter 3

Direction-sensitive responses to head turns in human superior temporal sulcus

3.1 Introduction

Humans and other primates share a remarkable ability to perceive where other individuals are looking and use this information to change their own attentional states (Deaner & Platt, 2003). Many higher order social cognitive processes depend on such gaze following behaviours (Frith & Frith, 2008; Klein, Shepherd & Platt, 2009). Although changes to gaze direction and head view are inherently dynamic, to date, the majority of human neuroimaging research has used static facial stimuli to study the neural representation of such social cues (Nummenmaa & Calder, 2009). In view of macaque electrophysiology evidence that neurons responsive to dynamic head turns do not respond to static views of the same head (Hasselmo et al., 1989), it is vital to explore the neural coding of dynamic social stimuli. Here, I demonstrate that a region in STS/STG contains a distributed representation of perceived head turn direction, thus supplying a necessary perceptual component to support a range of social behaviours.

Neurons in macaque anterior STS play a well-established role in representing the perceived direction of others' social attention cues, as conveyed by head orientation, gaze direction and body posture (Perrett et al., 1982, 1992, 1985; Wachsmuth et al.,

1994). However, these constitute only a minority of visually responsive STS neurons and are either spatially distributed (Hasselmo et al., 1989), or organised into fine-grained patches well beyond the resolution of conventional fMRI (Perrett et al., 1984). This distributed representation poses a significant signal-to-noise challenge for attempts to study similar effects with human fMRI, where each voxel likely samples millions of neurons in ways that are only indirectly related to the neuronal spike trains commonly measured in macaque electrophysiology (Logothetis, 2008; Kriegeskorte et al., 2009).

Unlike the typical anterior STS region identified by research in the macaque, most human fMRI studies report that social attention cues activate posterior STS and regions of adjacent STG and middle temporal gyrus (MTG) (Hein & Knight, 2008; Nummenmaa & Calder, 2009). Similar lateral posterior temporal regions are also more responsive to full-face views of faces than to control stimuli (Kanwisher et al., 1997; Andrews & Ewbank, 2004; Fox, Iaria & Barton, 2009). Most of these studies find that posterior STS is more responsive to averted than to direct gaze (Nummenmaa & Calder, 2009) but the opposite pattern has also been observed (e.g., Pageler et al., 2003; Pelphrey et al., 2004). Furthermore, posterior STS responds more when an actor gazes away from a target than when the gaze direction is congruent with the target location (Pelphrey et al., 2003), suggesting that posterior STS is influenced by contextual effects, rather than by the direction of the social attention cue as such. Even in the absence of overt contextual manipulations, comparisons between direct and averted gaze may indirectly manipulate the engagement of approach/avoidance mechanisms and other higher order social cognitive functions associated with direct and averted gaze, such as theory of mind responses to eye contact (Emery, 2000; Senju & Johnson, 2009; Shepherd, 2010). Thus, the litmus test for direction sensitivity is whether brain responses to different averted social attention cues can be distinguished in the absence of other contextual manipulations.

When such tests for direction sensitivity between different averted cues were carried out, one study found direction-sensitive fMRI adaptation to static images of gaze cues in right anterior, rather than posterior, STS (Calder et al., 2007). Another study that applied MVPA to a posterior STS ROI observed no distinction between different averted views of static heads (Natu et al., 2010) but did find that this ROI distinguished direct from averted head views across different head identities, suggesting an identity-

invariant representation. These head view effects are consistent with the pattern of univariate sensitivity for direct against averted gaze observed in previous univariate research (Nummenmaa & Calder, 2009). Considered collectively, this literature suggests a broad role for posterior STS in representing social attention cues but unlike the evidence from macaque anterior STS, there is little indication that posterior STS represents such cues in a direction-sensitive manner.

Outside the laboratory, cues to another's focus of attention are intrinsically dynamic in nature but this issue has received limited attention in controlled experiments. There is initial evidence that a small subset of neurons in macaque anterior STS are tuned to dynamic changes in head turn direction (Perrett et al., 1985; Hasselmo et al., 1989) but it remains unclear how the human brain codes such stimuli. In humans, posterior STS responds more to dynamic head turns than to both scrambled controls and static head views (Lee et al., 2010). However, neither anterior nor posterior STS has been found to show direction-sensitive coding of head turn direction, as measured by standard univariate fMRI (Lee et al., 2010). This absence of direction sensitivity is unsurprising, since neurons with such responses are unlikely to be clustered at a sufficiently large spatial scale to be detectable with univariate fMRI (Perrett et al., 1984; Hasselmo et al., 1989).

Multivariate pattern analysis has recently been applied to detect representations thought to be coded in fine-grained patterns beyond the resolution of standard fMRI (Kamitani & Tong, 2005; Haynes & Rees, 2006; Shmuel et al., 2010). In this chapter, I use MVPA to determine whether distributed response patterns in the human STS region contain distinct direction-sensitive codes for observed head turns. If a classifier can use response patterns from the STS region to distinguish between leftward and rightward head turns, this would suggest that the underlying response patterns code head turn direction. However, leftward and rightward motion can also produce classification effects in regions without selectivity for social attention cues (Kamitani & Tong, 2006). In order to avoid such confounding contributions of low-level motion, I included a set of rotating ellipsoid control videos. Previous work investigating head turn responses in macaque electrophysiology (Perrett et al., 1985; Hasselmo et al., 1989) or direction-specific responses to static gaze (Calder et al., 2007) did not include such non-social controls, so an important aim of the current study was to establish that any direction-sensitive effects are specific to social stimuli. Furthermore, I aimed

to localise pattern effects to specific regions through the use of a searchlight algorithm that operated within the anatomically defined STS region. The STS region in this study included STG and MTG, in line with previous findings that social perception and gaze stimuli produce peaks that sometimes fall outside the STS proper (Allison, Puce & Mccarthy, 2000; Nummenmaa & Calder, 2009).[1]

3.2 Experiment 1: Direction-sensitivity to head turns

3.2.1 Materials and methods

3.2.1.1 Participants

Twenty-one right-handed healthy volunteers with normal or corrected to normal vision participated in the study (12 males, mean age 29 years, age range 22–38). Volunteers provided informed consent as part of a protocol approved by the Cambridge Psychology Research Ethics Committee. Four volunteers were removed from further analysis: two due to poor performance at the behavioural task whilst in the scanner (accuracy of less than 50%) and two due to fatigue and excessive head movements.

3.2.1.2 Experimental design

Volunteers viewed 1000-ms video clips of 45° leftward and rightward head turns and comparable ellipsoid rotations (Figure 3.1). Volunteers were instructed to monitor the stimulus set for infrequent deviant response trials (1 of the 8 experimental videos, rotated 4° from the upright position) and responded to detected deviants with a button press. The deviant response trials were drawn from all experimental conditions and the degree of rotation was chosen after behavioural pilot tests to produce an attentionally demanding task without ceiling effects.

Two actors with matched head motion patterns were selected for the head turn videos. The ellipsoid control stimuli were rendered and animated in Matlab (Mathworks) and were texture mapped with the Fourier-scrambled face textures from the two head identities. The two motion directions were created by mirror reversing video

[1]The research described here was published in *Cerebral Cortex* (Carlin, Rowe, Kriegeskorte, Thompson & Calder, 2011). It is reprinted here in chapter form (see Appendix A for journal form).

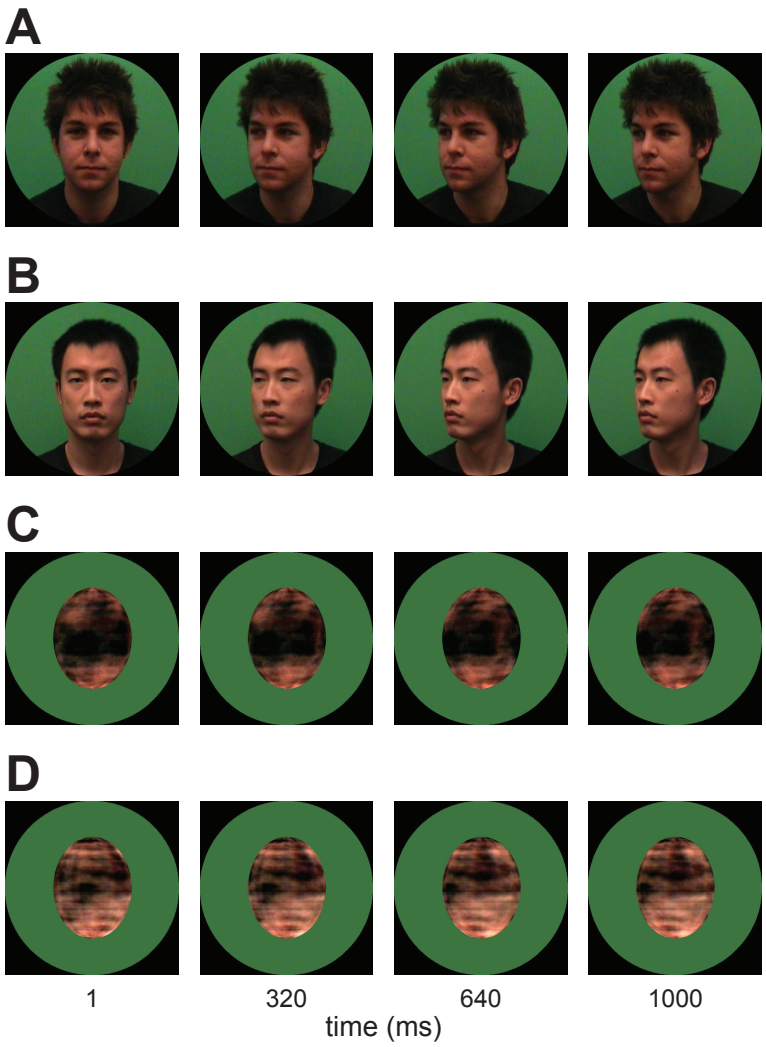


Figure 3.1: Example video frames for turning heads (**A-B**) and rotating ellipsoids (**C-D**). The videos were presented at 24 frames per second. All video frames are from leftward motion conditions. Rightward conditions were created through mirror reversal of the same video clips. The two ellipsoid identities (C-D) were created by Fourier-scrambling face textures from the first frame of the two head videos (A-B).

clips with a single direction, thus ensuring that the stimulus set was physically matched across motion directions. This produced a total of 8 stimuli (two heads, two ellipsoids, each rotating leftward or rightward), which were treated as individual conditions.

The stimuli were back-projected onto a screen in the scanner, which volunteers viewed via a tilted mirror. The stimuli were presented on a black background within a circular aperture (7° visual angle in diameter). The experiment was controlled using Matlab and the Psychophysics toolbox (Brainard, 1997).

The experiment was divided into sets of 240 trials, each of which was independently randomised. Parameter estimates from each set formed an independent set of training examples for classification. The trials were presented within a rapid event-related design. Four volunteers completed a 6-set version of the experiment (approximately 40 min effective time) and 13 completed a 12-set version (80 min). Each set contained 240 trials: 80 null trials, where a fixation cross remained on the screen throughout the trial (1500 ms) and 160 experimental trials (80 heads, 80 ellipsoids), where each trial consisted of a video clip (1000 ms) followed by fixation (500 ms). Each condition was repeated 18 times in a set. Sixteen deviant response trials were randomly sampled from the experimental conditions and responses to these trials were modelled with a separate nuisance regressor of no interest. The trials within the set were presented in a pseudo-randomised order, where repeats of the same trial were slightly clustered in order to increase design efficiency (Henson, 2003). Every second set was followed by a 15 s rest period, which was cued by a text prompt on the screen. The scan acquisition continued during the rest periods and volunteers were instructed to remain still.

3.2.1.3 Imaging acquisition

Scanning was carried out at the MRC Cognition and Brain Sciences Unit, Cambridge, United Kingdom, using a 3-T TIM Trio magnetic resonance imaging (MRI) scanner (Siemens), with a head coil gradient set. Functional data were collected using high-resolution echo planar T2*-weighted imaging (EPI, 40 oblique axial slices, time to repeat (TR) 2490 ms, time to echo (TE) 30 ms, in-plane resolution 2×2 mm, slice thickness 2 mm $\pm 25\%$ slice gap, 192×192 mm field of view). The acquisition window was tilted up approximately 30° from the horizontal plane to provide complete

coverage of the occipital and temporal lobes. Preliminary pilot tests suggested that the use of this high-resolution EPI sequence resulted in reduced signal dropout in the anterior STS region, relative to a standard resolution sequence ($3 \times 3 \times 3.75$ mm voxels). All volumes were collected in a single continuous run for each volunteer. The initial 6 volumes from each run were discarded to allow for T_1 equilibration effects. T_1-weighted structural images were also acquired (magnetisation prepared rapid gradient echo (MPRAGE), 1 mm isotropic voxels).

3.2.1.4 Imaging analysis

Imaging data were processed using statistical parametric mapping (SPM) 5 (www.fil.ion.ucl.ac.uk/spm). All functional volumes were realigned to the first non-discarded volume, slice time corrected and co-registered to the T_1 structural volume. The processing pathways for univariate analysis and MVPA diverged after these common steps (Figure 3.2).

Univariate analysis was carried out using standard processing steps in SPM5. Structural volumes were segmented into grey and white matter partitions and spatially normalised to the Montreal Neurological Institute (MNI) template using combined segmentation and normalisation routines. Functional volumes were normalised according to the parameters of this transformation, smoothed (10 mm FWHM Gaussian kernel) and high pass filtered to remove low frequency drift (128 s cut-off period).

Subject-specific generalised linear models were used to analyse the data. The models included one regressor per condition and nuisance regressors for deviant response trials, volunteer responses to non-deviant trials and for nulling scans that contained excessive noise or movement (Lemieux, Salekhaddadi, Lund, Laufs & Carmichael, 2007; Rowe, Eckstein, Braver & Owen, 2008, greater than 10 units intensity difference from the mean scaled image variance or more than 0.3 mm translational or 0.035 radians rotational movement relative to the previous volume). The volunteer-specific models included $0-135$ such scan nulling regressors (mean 35). The experimental predictors were convolved with a canonical haemodynamic response function and contrast images were generated based on the fitted responses. These contrast images were then entered into second-level permutation-based random effects models using SnPM (Nichols & Holmes, 2001, 10000 permutations, 10 mm FWHM variance smoothing).

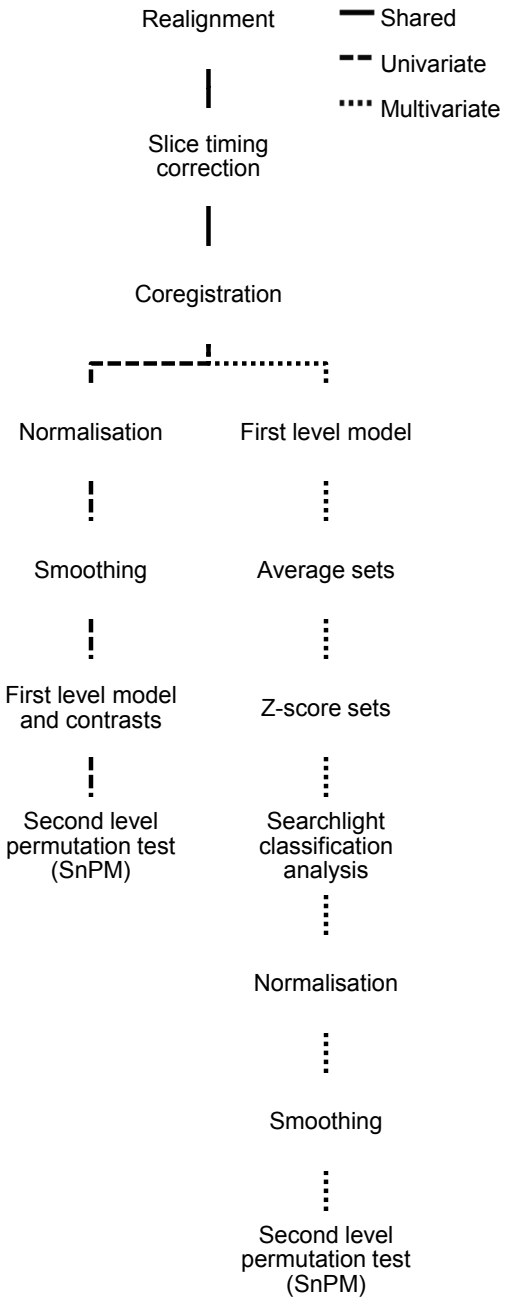


Figure 3.2: Processing pathways for the fMRI analysis in Experiment 1. All processing nodes take the result of the previous node in the hierarchy as input. With the exception of the searchlight classification analysis, all processing steps were implemented using standard SPM 5 functionality.

Multivariate pattern analyses were carried out using functional volumes that had been realigned and slice time corrected but had not been spatially normalised to the MNI template (Figure 3.2). Each volunteer's data were modelled using a generalised linear model with similar regressors as in the univariate analysis with the exception that each set of trials was modelled using a separate set of regressors. Individual parameter volumes from the first half of the data set were then averaged pairwise with the corresponding volume from the second half of the data set, thus reducing session effects at the expense of halving the number of training examples. This produced 3 or 6 final sets of examples to be used for classification, depending on the number of available sets before averaging. The example volumes were Z-scored so that each voxel within a set had a mean of 0 and a standard deviation of 1 across examples in that set. Finally, each example was grey matter masked using the tissue probability maps generated by the segmentation processing stage.

The resulting example volumes were used as input to a linear SVM classifier (as implemented in PyMVPA, Hanke et al., 2009). All MVPA used a searchlight algorithm (Kriegeskorte et al., 2006), in which classification is carried out within a spherical region (5 mm radius) that is moved through the volume. Leave-one-out cross-validated classification accuracy estimates (percent correct) were mapped back to the centre of each searchlight to produce a classification accuracy map.

The classification accuracy maps for each volunteer were normalised to MNI space, smoothed (10 mm FWHM) and entered into second-level nonparametric random effects models in SnPM. I used nonparametric tests because the discontinuous nature of the grey matter masked data means that conventional FWE correction for multiple comparisons using random field theory in SPM5 would be inappropriate.

In line with the hypothesised site of the effects, I restricted the primary analysis to the right STS region, which was defined anatomically based on the mean T_1 volume for the sample. In line with previous evidence that social perception and eye gaze effects in the STS region extend into STG and MTG (Allison et al., 2000; Nummenmaa & Calder, 2009), the mask included these gyri, whilst leaving out voxels in inferior temporal sulcus (inferior) or lateral fissure (superior, Figure 3.3).

I report p values corrected for multiple comparisons (FWE, $p < 0.05$) within this ROI (5162 voxels, y -58 to 22 mm MNI). I also carried out an exploratory analysis in a mirror-reversed version of the STS mask to test for effects in left STS. The use of a

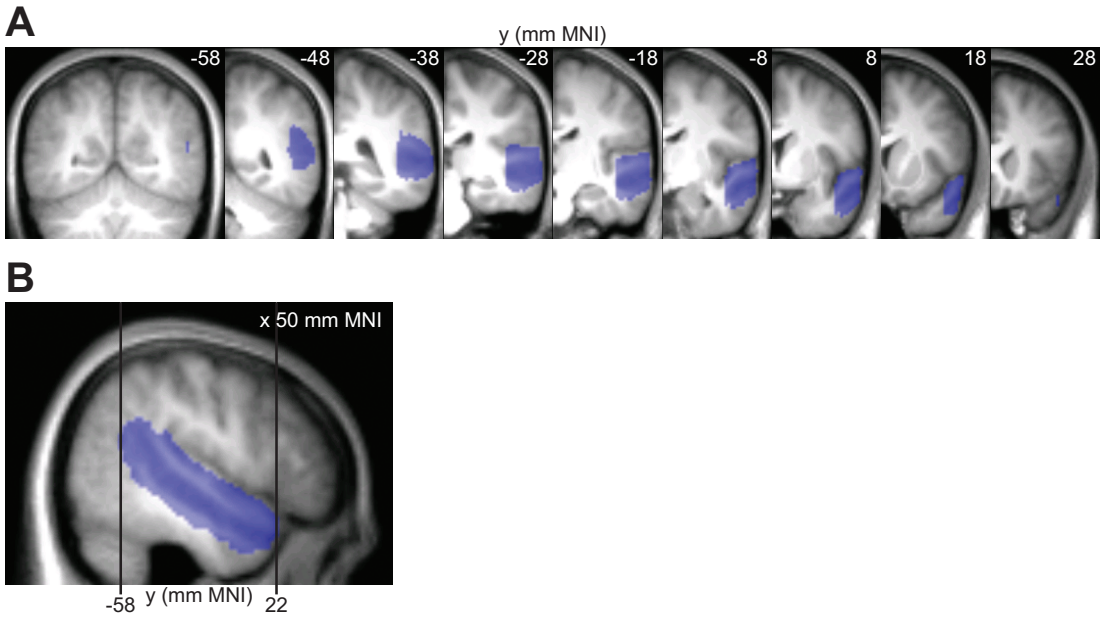


Figure 3.3: (**A**) Coronal sections of the anatomical mask for the right STS region. (**B**) Saggital section showing the full length of the mask in the y plane.

mirror-reversed mask sacrifices some anatomical precision in left STS but preserves the same voxel count and spatial structure in both masks. Visual inspection of the relation between the left STS mask and the mean T_1 volume suggested that the mask followed the anatomy of the sulcus in a comparable manner to the right STS mask. Finally, effects that survived correction for the full volume are also reported (FWE, $p < 0.05$). All analyses were restricted to a group grey matter mask, which was formed by the union of each volunteer's normalised individual grey matter mask. This mask ensured that I only considered effects in regions actually covered by the searchlight analysis.

3.2.2 Results

3.2.2.1 Behavioural task

Volunteers were asked to detect the occasional 4° rotation of the video stimuli and were able to detect such deviant response trials adequately (mean accuracy 71%, standard error 4%). A repeated measures analysis of variance (ANOVA) of accuracy scores with the factors of stimulus type (head, ellipsoid) and motion direction (leftward, rightward)

yielded no main effects and no interaction ($F_{(1,16)} < 2.4$, $p > 0.14$ for all effects), suggesting that volunteers did not assign attention differently to the heads and ellipsoids or to the two motion directions.

3.2.2.2 Multivariate pattern analysis: STS

My primary hypothesis was that the right anterior STS region distinguishes between leftward and rightward perceived head turns. In line with this prediction, a group analysis of the MVPA searchlight results for the right STS region showed that classification of head turn direction was significantly more accurate than expected by chance in a right anterior STS/STG site ($p = 0.005$ FWE, $[50, 4, -14]$ mm MNI, Figure 3.4A; for individual subject results, see Figure 3.5). By comparison, left-right classification of rotation direction in the ellipsoid control stimuli exceeded chance in middle STS ($p = 0.037$ FWE, $[50, -14, -10]$ mm MNI, Figure 3.4B).

The peaks of these head turn and ellipsoid rotation effects were approximately 18 mm apart and the activated regions did not overlap, which raises the question of how distinct the two effects are. I addressed this by computing the difference between the classification maps for head turn and ellipsoid rotation in each volunteer. These difference maps were entered into a group analysis, which showed that left-right classification was more accurate for head turns than for ellipsoid rotations in right anterior STS/STG ($p = 0.027$ FWE, $[52, 12, -12]$ mm MNI, Figure 3.4C). This effect overlapped with the head turn classification effect (8 mm distance between peaks, 40% overlap), suggesting a common origin. No STS region showed significantly more accurate direction classification for ellipsoid rotations than for head turns.

I tested whether the left-right head turn codes were invariant to head identity by training the classifier on the left-right turns of one head and applying the learned weights to left-right turns of the other head. Left-right classification did not generalise across head identity at any site in right STS. Similarly, there was no significant left-right generalisation across ellipsoid identities and no left-right generalisation across stimulus type (head and ellipsoid).

I also carried out an exploratory analysis of effects in the left anatomically defined STS region. No left STS regions showed above-chance classification of observed head turn direction. However, a region in left anterior STS distinguished ellipsoid rota-

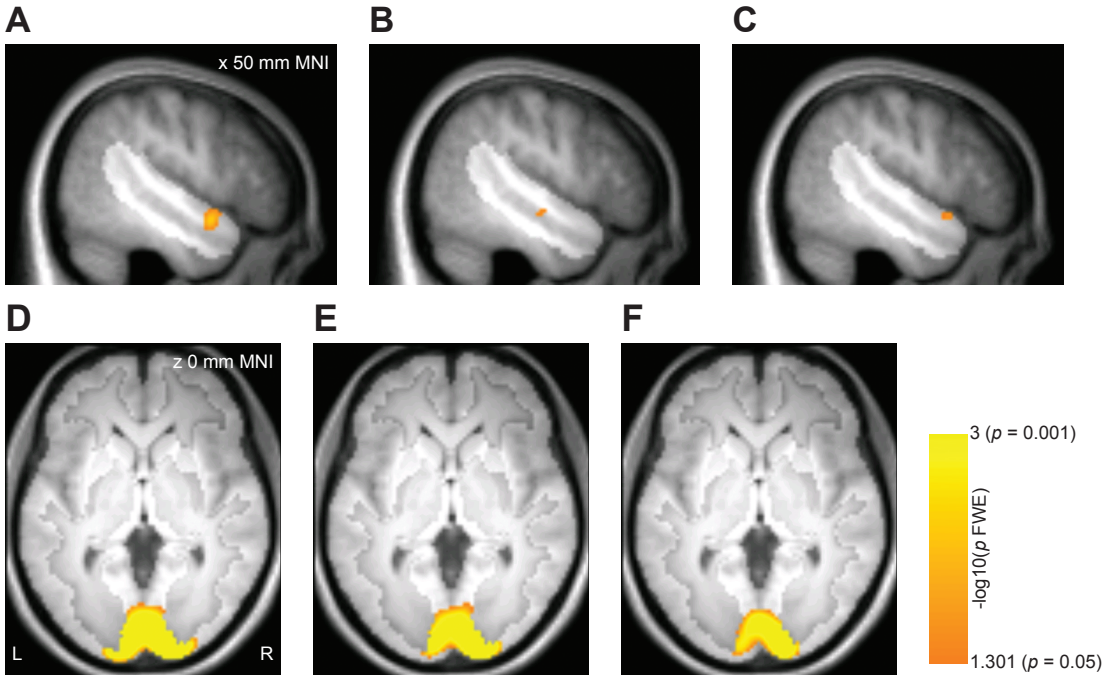


Figure 3.4: Group results for MVPA, displayed on the mean T_1 volume for the sample. Effects are displayed corrected for multiple comparisons within the right STS region (panels A–C; hypothesis-driven analysis, $p < 0.05$ FWE) or the full grey matter volume (panels D–F; exploratory analysis, $p < 0.05$). The highlighted portion of each panel shows the extent of the mask. (**A**) Classification of left-right head turns in the right STS/STG region. (**B**) Classification of left-right ellipsoid rotations in the right STS region. (**C**) Right STS regions where left-right classification of head turns was more accurate than classification of ellipsoid rotations. (**D**) Classification of left-right head turns in the full grey matter volume. (**E**): Grey matter regions where left-right classification of head turns was more accurate than classification of ellipsoid rotations. (**F**) Grey matter regions where the weights acquired by training the classifier on left-right head turns for one head identity generalised to left-right head turns in the other head identity.

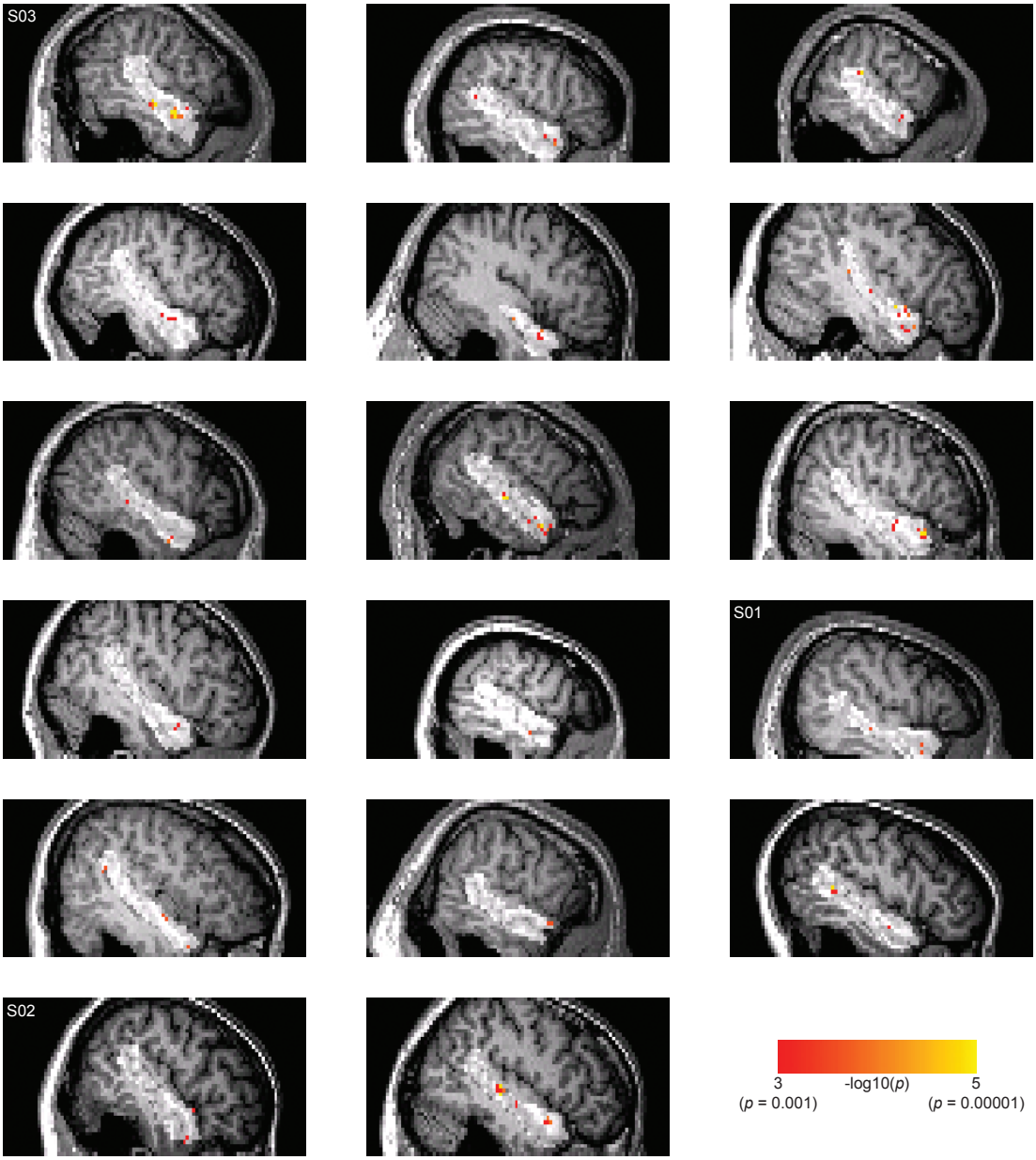


Figure 3.5: Single volunteer searchlight results for the left-right head turn classification effect. Individual searchlight maps are thresholded at $p < 0.001$ (uncorrected, binomial test) and are masked to include only searchlight centres falling inside the right STS anatomical mask (highlighted). Codes for volunteers who were used in the final analysis of the follow-up scans are shown in the relevant panels. It can be seen that anterior STS effects are apparent in most volunteers. However, note that single subject results were normalised and smoothed (10 mm FWHM) prior to the group analyses reported in the main text (see Section 3.2.1.4) and thus any direct comparison between group and single volunteer results should be made with caution.

tion direction with above-chance accuracy ($p = 0.01$ FWE, $[-56, -8, -16]$ mm MNI, Figure 3.6A). Direction classification accuracy was significantly higher for ellipsoid rotation than for head turns in a similar region ($p = 0.041$ FWE, $[-58, -4, -16]$ mm MNI, Figure 3.6B). No other classification effects were significant in this ROI.

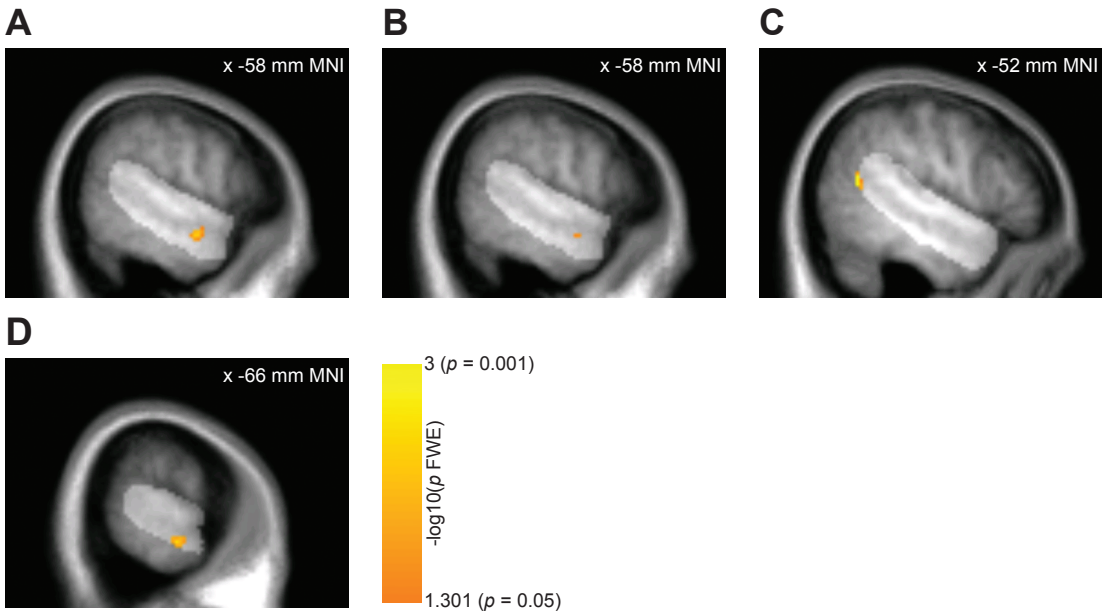


Figure 3.6: Group results for the left STS anatomical ROI (highlighted, $p < 0.05$ FWE-corrected within this region), displayed on the mean T_1 volume for the sample. **A**: Classification of left-right ellipsoid rotation. **B**: Regions where left-right classification of ellipsoid rotations was more accurate than classification of head turns. **C**: Region with greater univariate responses to heads than to ellipsoids. **D**: Regions with greater univariate responses to ellipsoids than to heads.

3.2.2.3 Multivariate pattern analysis: Whole-brain

Beyond my hypothesis-driven search within the anatomically defined right STS region, I also carried out an exploratory analysis within the full grey matter-masked volume to identify other effects of interest. Classification of left-right head turns exceeded chance in a region including calcarine sulcus and occipital pole ($p < 0.001$ FWE, $[16, -96, 0]$ mm MNI, Figure 3.4D). This region is likely to include visual areas V1, V2 and V3, but in the absence of a retinotopic localiser I use the general term early visual cortex

to describe this region. Left-right ellipsoid classification did not produce significant effects in any region. Left-right classification was significantly more accurate for head turns than for ellipsoid rotations in a similar early visual region ($p < 0.001$ FWE, $[14, -96, 2]$ mm MNI, Figure 3.4E). A similar region in early visual cortex also allowed left-right classification to generalise across head identities ($p < 0.001$ FWE, $[14, -96, 2]$ mm MNI, Figure 3.4F) but not across stimulus types. No regions outside of early visual cortex showed significant effects for any of these comparisons.

3.2.2.4 Univariate analysis: STS

I used a univariate analysis to address whether the observed classification effects could be attributed to large-scale response level differences between the conditions. To make comparisons between MVPA and univariate results simpler, the univariate analysis also used nonparametric permutation-based random effects analysis of group effects (SnPM, for details, see Section 3.2.1.3). I also explored whether direction classification of head turns co-localised with greater univariate responses to heads than to ellipsoids.

No regions inside the anatomically defined right STS ROI responded selectively to one head turn direction over the other or to one ellipsoid rotation direction over the other, suggesting that the left-right classification effects in this region did not co-occur with large-scale univariate direction sensitivity.

Collapsing across motion direction, right posterior STS responded significantly more to heads than to ellipsoids ($p = 0.002$ FWE, $[48, -44, 16]$ mm MNI, Figure 3.7A), while a region in middle STG bordering on the edge of the ROI responded more to ellipsoids than to heads ($p = 0.004$ FWE, $[60, 0, 0]$ mm MNI, Figure 3.7B). Thus, univariate selectivity for heads over ellipsoids occurred in posterior STS, 57 mm from the left-right head turn classification peak in anterior STS/STG. The peaks for univariate selectivity for ellipsoids over heads and for left-right ellipsoid rotation classification were separated by 20 mm. Neither of the univariate effects overlapped with the classification effects.

Within the left STS ROI, a posterior region responded more to heads than to ellipsoids ($p = 0.004$ FWE, $[-52, -58, 14]$ mm MNI, Figure 3.6C) and left middle STS responded more to ellipsoids than to heads ($p = 0.014$, $[-66, -18, -14]$ mm MNI, Figure 3.6D), mirroring the results obtained in the right STS region. No left STS

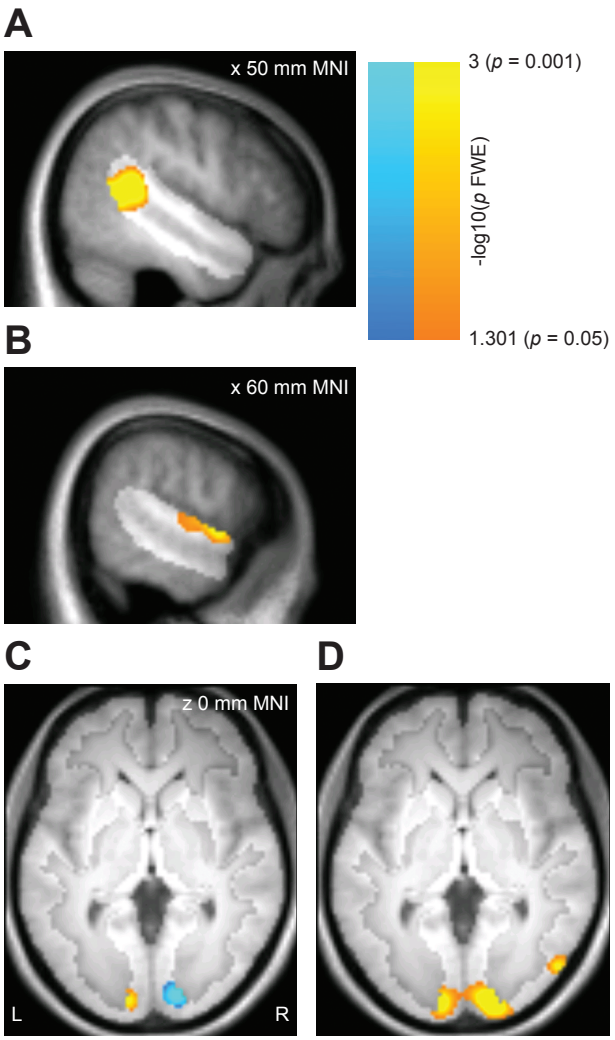


Figure 3.7: Group results for the univariate analysis, displayed on the mean T_1 volume for the sample. Effects are displayed corrected for multiple comparisons within the right STS region (panels A–B; hypothesis-driven analysis, $p < 0.05$ FWE) or the full grey matter volume (panels C–D; exploratory analysis, $p < 0.05$ FWE). The highlighted portion of each panel shows the extent of the mask. (**A**) Greater univariate responses to heads than to ellipsoids in the right STS region. (**B**) Greater univariate responses to ellipsoids than to heads in the right STS region. (**C**) Grey matter regions with greater univariate responses to left than to right head turns (warm colours) or with greater univariate responses to right than to left head turns (cool colours). The effects do not overlap at any voxel. (**D**) Grey matter regions with greater univariate responses to heads than to ellipsoids.

regions responded preferentially to head turn or ellipsoid rotation in one direction relative to another. No other comparisons reported above were significant in the left STS analysis.

3.2.2.5 Univariate analysis: Whole-brain

A univariate analysis of the grey matter-masked full volume revealed significant univariate selectivity for left over right head turns that was restricted to left early visual cortex ($p < 0.001$ FWE, $[-12, -94, 0]$ mm MNI) and conversely, selectivity for right over left head turns restricted to right early visual cortex ($p < 0.001$ FWE, $[14, -92, 4]$ mm MNI, Figure 3.7C). These effects almost completely overlapped the left-right head turn classification effect in early visual cortex (100% overlap for left over right, 91% overlap for right over left), suggesting that the classification effects co-occurred with large-scale univariate effects. Note that the laterality of these early visual effects is opposite to what would be expected for a stimulus that moves into the right and left visual hemifields, a point I return to below. No regions showed a preference for one ellipsoid rotation direction over the other in the whole-brain analysis.

A comparison of univariate responses to heads over ellipsoids and ellipsoids over heads revealed a network of activations (Table 3.1). Of primary interest to the current study, bilateral early visual cortex responded more to heads than to ellipsoids ($p < 0.001$ FWE, $[18, -96, -4]$ mm MNI, Figure 3.7D) and this early visual effect overlapped the left-right head turn classification effect (91% overlap). Thus, the left-right head turn classification effects occurred in a region where I also observed univariate selectivity for head turn direction and preferential responses to heads over ellipsoids. Bilateral regions in posterior MTG also responded more to heads than to ellipsoids (right: $p = 0.001$ FWE, $[52, -74, 2]$ mm MNI. Left: $p = 0.001$ FWE, $[-50, -72, 14]$ mm MNI). These coordinates are close to those previously reported for motion area MT (Dumoulin, Bittar, Kabani, Baker, Le Goualher, Pike & Evans, 2000) (conversion from Talairach to MNI coordinates with tools from Evans, Zilles, Lancaster, Martinez, Mazziotta, Fox, Tordesillas-Gutierrez & Salinas, 2007). Because I did not include a specific localiser scan to distinguish MT from other motion areas, I refer to this region as MT+. The MT+ regions showed no direction-sensitive responses in the univariate or classification analyses, even at reduced thresholds ($p < 0.01$, un-

corrected).

Table 3.1: Regions with significant response level differences between heads and ellipsoids in the univariate analysis ($p < 0.05$, FWE corrected for grey matter regions in the full volume).

Region	Hemisphere	p	Peak (mm MNI) x	y	z
Heads > Ellipsoids					
Early visual cortex	L/R	<0.001	18	-96	-4
Fusiform gyrus	R	<0.001	40	-44	-20
Middle temporal gyrus (MT+)	L	<0.001	-50	-72	14
Superior temporal sulcus	R	0.001	48	-44	16
Middle temporal gyrus (MT+)	R	0.001	52	-74	2
Middle frontal gyrus	R	0.016	46	4	52
Ellipsoids > Heads					
Parahippocampal gyrus	L	<0.001	-26	-52	-16
Parahippocampal gyrus	R	<0.001	28	-42	-12
Middle occipital gyrus	L	<0.001	-32	-90	10
Middle occipital gyrus	R	<0.001	36	-84	6
Anterior cingulate	L	0.007	-8	46	2
Lateral sulcus	R	0.007	58	2	2

3.3 Interim discussion

The pattern of univariate effects in early visual cortex suggested the presence of eye movements in the experiment. If volunteers tracked the heads as they turned, this would have placed the stimulus primarily in the hemifield ipsilateral to the direction of motion, which could explain the ipsilateral univariate activations in early visual cortex. Eye tracking was not available when Experiment 1 was undertaken, so I carried out follow-up eye tracking and fMRI experiments with 3 principal aims: first, to test whether the head turns used in Experiment 1 elicit eye movements; second, to assess whether the eye movement effects could be removed with a revised experimental paradigm; finally, to test whether the fMRI effects reported in the main text remained in the absence of statistically significant eye movement effects.

3.4 Materials and methods: Experiments 2-4

Five volunteers from the final sample used in Experiment 1 returned to participate in additional behavioural and fMRI experiments. Eye movements were monitored using a video-based infra-red eye tracker (500 Hz acquisition outside the scanner, 50 Hz acquisition inside the scanner; Sensomotoric Instruments). I analysed the change in horizontal fixation position at the end relative to the start of each trial using custom code developed in Matlab.

Imaging data were acquired and analysed using identical parameters as in Experiment 1 with the exception that no averaging of the first and second half of the experiment was carried out since this would have yielded an unacceptably small number of observations for first-level statistics. Furthermore, each set was scanned in a separate run to allow recalibration of the eye tracker between sets. As in Experiment 1, I used a searchlight analysis. I based single-volunteer inference on binomial tests at each voxel in the ROI.

3.5 Experiment 2: Eye tracking with the original design

Five volunteers carried out an abbreviated version of Experiment 1 outside the scanner (3 sets, 540 trials) while their eye position was monitored. First-level ANOVAs revealed that each volunteer showed a significant stimulus type (head, ellipsoid) by motion direction (leftward, rightward) interaction (Table 3.2). This interaction reflected consistent fixation shifts in the direction of the head turns with non-significant or weaker fixation shifts in the direction of the ellipsoid rotations.

3.6 Experiment 3: Eye tracking with revised design

I carried out a second eye tracking experiment with a revised paradigm that included a fixation cross during the presentation of the video clips. Volunteers were also strongly instructed to maintain fixation at all times. I included only the head turn conditions in order to obtain a maximal number of trials for the head left-right comparison whilst

Table 3.2: Horizontal fixation change analysed using volunteer-specific ANOVAs with the factors of stimulus type (head, ellipsoid) and direction (leftward, rightward).

		F		
Volunteer	df	Stimulus type	Direction	Interaction
S01	1, 417	11.79*	6.48*	10.14*
S02	1, 382	2.29	17.92*	21.11*
S03	1, 418	8.36*	26.8*	31.75*
S04	1, 408	0.13	14.54*	51.16*
S05	1, 390	1.39	34.09*	76.72*

* $p < 0.05$

minimising volunteer fatigue. In this second experiment, the head turn left-right eye movement effect was reduced to non-significance in 4 of 5 volunteers (Figure 3.8A-C).

3.7 Experiment 4: fMRI experiment with revised design

I tested whether the main classification findings in STS/STG and early visual cortex survived in the absence of eye movements by carrying out a second fMRI experiment with the revised experimental paradigm from Experiment 3. I recruited the 4 volunteers who showed no significant eye movement effects in Experiment 3. Volunteers completed a full 6-set version of Experiment 3 (1080 trials, for details, see Section 3.2.1), while their eye position was monitored. One of the 4 scanned volunteers showed a significant fixation shift in response to the head turns whilst being scanned ($F_{(1,390)} = 8.72$, $p = 0.003$). This volunteer was removed from further analysis.

Although the 3 remaining volunteers showed no significant eye movement effects (as observed in separate tests before and during scanning), left-right classification of head turns in the right anterior STS region was greater than chance in two volunteers ($p < 0.05$, Bonferroni FWE corrected for the right STS mask) and at reduced thresholds in the third ($p < 0.001$, uncorrected, 3.8D-F). The final volunteer also showed an effect in posterior STS ($p < 0.05$, FWE).

All 3 volunteers showed significant left-right head turn classification effects in early

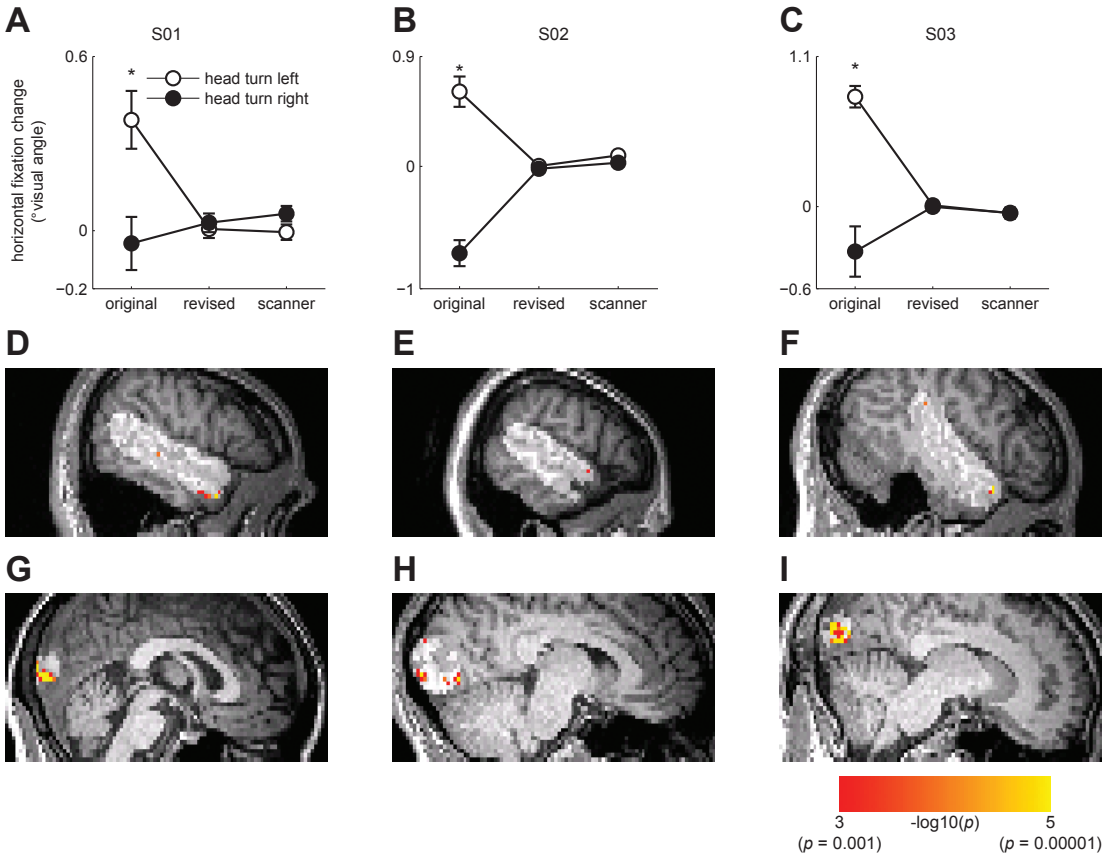


Figure 3.8: Follow-up eye tracking and fMRI experiments (Experiments 2-4). (**A–C**) Mean horizontal fixation change plotted separately for the 3 volunteers selected for the final analysis in the revised fMRI experiment. Positive values reflect a leftward shift in fixation over the trial, while negative values reflect a rightward shift. The horizontal axis gives fixation performance in the original task (Experiment 2), the revised task (Experiment 3) and the revised task as measured during fMRI (Experiment 4). The error bars give ± 1 standard error of the mean. Comparisons with significant differences between the head turn directions are highlighted by asterisks (t tests, $p < 0.05$). It can be seen that the revised design abolished the eye movement effect in these volunteers. (**D-F**) Left-right head turn classification results for the 3 volunteers in the final sample of Experiment 4. The volunteers are shown in the same order as in A-C. Results are overlaid on each volunteer's T_1 volume and are masked to only include effects within the highlighted right STS region ($p < 0.001$, uncorrected). It can be seen that even in the absence of eye movement effects, anterior STS/STG codes head turn direction. (**G-I**) Results as in D-F but masked to show effects within a 20 mm radius of the peak early visual head turn classification effect from the main study. It can be seen that the effects in early visual cortex also remain when eye movements are controlled.

visual cortex ($p < 0.05$ Bonferroni FWE corrected for a 20 mm radius sphere centred on the peak head turn classification effect in Experiment 1, Figure 3.8G-I). However, unlike Experiment 1, where this effect was joined by univariate response preferences for head turns in a direction ipsilateral to the visual hemifield (3.7A), I now observed preferentially contralateral responses to head turn direction ($p < 0.001$, uncorrected, Figure 3.9B). Thus, although the classification effects in early visual cortex were also accompanied by univariate effects in the revised Experiment 4, the laterality of these univariate effects was reversed.

3.8 General discussion

Appropriate social behaviour is dependent on accurately inferring where others are attending. In the visual domain, this inferential process is likely to involve direction-sensitive coding of social attention cues, such as head turns. In experiments, these stimuli are often abstracted to static views, which fails to capture their dynamic character in natural social interaction. Here, I demonstrate that response patterns in human right anterior STS/STG distinguish between leftward and rightward dynamic head turns. Furthermore, left-right head turns were significantly more discriminable in this region than left-right ellipsoid control stimuli. A similar analysis of the left STS region revealed no left-right classification of head turn direction at any voxel in the ROI.

3.8.1 Relation to macaque electrophysiology and previous fMRI studies

The peak coordinates for left-right classification of head turn direction in the current study are in close proximity to a previous demonstration of direction-sensitive fMRI adaptation to static gaze (Calder et al., 2007, 16 mm distance between peaks). Considered collectively, these results suggest a general role for right anterior STS/STG in supplying higher order social cognitive processes with information about the direction of another's attentional shifts, whether these are conveyed by static gaze in a front-facing head or dynamic head turns. Consistent with this social role, I also demonstrate that direction sensitivity does not extend to non-social control stimulus motion in this region. An important question is whether such direction-sensitive responses in humans

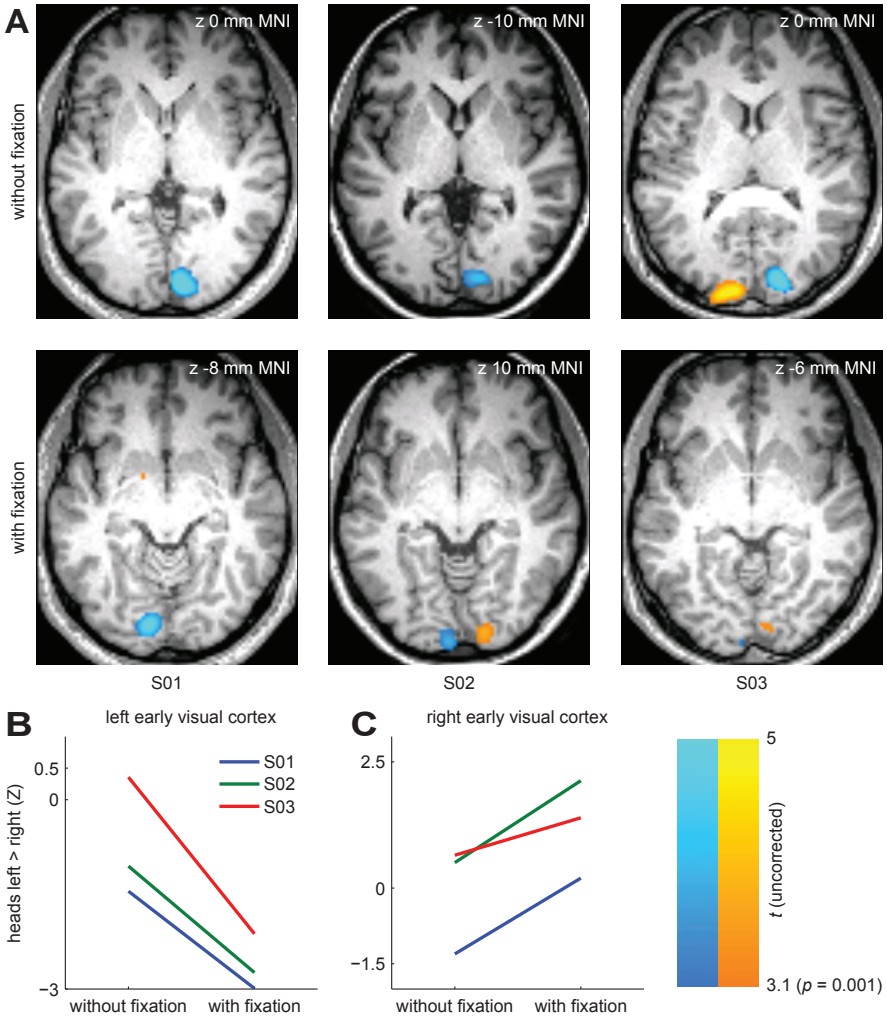


Figure 3.9: Univariate responses to head turn direction in the original and revised experiments. **A**: Regions responding more to left than to right head turns (warm colours) and regions responding more to right than to left head turns (cool colours). Results are shown for the original Experiment 1 (without fixation) and for the revised Experiment 4 (with fixation) in rows with the 3 volunteers from the final analysis of Experiment 4 in columns (S01, S02, S03). It can be seen that in Experiment 1 (without fixation), all selective responses in early visual cortex are ipsilateral to the direction of motion in the head stimuli (e.g. right early visual cortex responds more to rightward than to leftward head turns). In the follow-up Experiment 4 (with fixation), any selective responses in early visual cortex are instead contralateral to the direction of motion in the head stimuli. **B-C**: Selectivity for left against right head turns (Z statistics), plotted for the peak voxel in left and right early visual cortex, as selected by the contrast heads and ellipsoids against baseline. It can be seen that for all volunteers and hemispheres, the left-right head turn selectivity tends to become more contralateral in Experiment 4 (with fixation), as compared to Experiment 1 (without fixation).

to dynamic and static social cues are driven by a single representation of the direction of another's social attention (Perrett et al., 1992) or whether dynamic information is coded separately, as indicated by the finding that macaque STS neurons tuned to head turn motion do not respond to static head view displays (Perrett et al., 1985; Hasselmo et al., 1989).

Neurons in macaque anterior STS are tuned to the direction of social attention cues (Hasselmo et al., 1989; Perrett et al., 1992). However, most human fMRI studies have reported gaze or head turn effects in posterior rather than anterior STS regions (Nummenmaa & Calder, 2009). Our classification effects appear more consistent with the typical recording site in macaque anterior STS than with previous univariate fMRI effects in human posterior STS. Compared with standard univariate comparisons between experimental conditions, MVPA and fMRI adaptation techniques may confer greater sensitivity (Haynes & Rees, 2006). This increased sensitivity makes more rigorous comparisons possible, for instance between left and right averted social attention cues. Accordingly, I also observed greater consistency between human fMRI and single unit evidence from the macaque (see also Kamitani & Tong, 2005; Calder et al., 2007). Known human-macaque discrepancies in the function of posterior STS and surrounding areas suggest that a simple correspondence between human and macaque may not apply to all high-level visual areas (Orban et al., 2004), but such a simple correspondence nevertheless offers a useful working model for the representation of social attention cues.

3.8.2 Relation to univariate face selectivity

The pattern of results I observed in posterior and anterior portions of the right STS region also highlights how large-scale univariate response level differences can dissociate from multivariate classification performance (Haxby et al., 2001; Hanson et al., 2004; Hanson & Halchenko, 2008). Similar to previous studies (Andrews & Ewbank, 2004), I found that right posterior STS responded more to heads than to ellipsoids, while no such preferential responding was observed in anterior STS/STG. The left-right head turn classification effects showed the opposite pattern, with significant effects in anterior but not posterior regions. There are clear parallels between this pattern of effects and a recent report where face identity classification was possible in an an-

terior inferotemporal region, which did not respond preferentially to faces over places, while no such face identity effects appeared in the more posterior FFA, even though this region responded more to faces than to places (Kriegeskorte et al., 2007). Face identity and head turn direction are both important dimensions for face processing, yet multivariate sensitivity for manipulations along these dimensions does not appear to co-localise with univariate selectivity for faces over other object categories. Although more systematic studies of these within- and between-category dissociations are needed before their theoretical implications for face perception can be fully considered, the current results indicate that studies where data analysis is restricted to functional ROIs defined by face selectivity are at risk of missing potentially important effects (Haxby et al., 2001; Friston et al., 2006).

3.8.3 Effects of head identity

Neurons with social attention responses in macaque STS are often invariant to the identity of the individual conveying the cue (Perrett et al., 1992). In this study, I observed no generalisation between response patterns evoked by left-right head turns across the two identities. Although there is some initial evidence to suggest that STS neurons can code both head view and head identity (Perrett et al., 1984), it is, in my view, unlikely that the representation across STS is identity-specific. For instance, it has previously been shown that direct and averted static head views can be distinguished across identity in posterior STS (Natu et al., 2010). Given that separate training of left-right classification for each identity involves half as much data as compared with when this dimension is collapsed, it is more likely that Experiment 1 did not have sufficient power to detect any such identity-invariant head turn representations.

3.8.4 Eye movement effects

Experiment 2 suggested that volunteers' eye movements tended to follow the direction of head turns, thus presenting a potential confound to the interpretation of the fMRI results in Experiment 1. To rule out an eye movement account of the reported classification effects, I demonstrated in Experiment 4 that a subset of volunteers from Experiment 1 showed significant left-right head turn classification in the right STS region, even though these volunteers showed no significant eye movement effects during

pretests or whilst in the scanner. Thus, even though the results of Experiment 1 are potentially limited by an eye movement confound, the head turn direction codes in the right anterior STS region remain when this confound is removed (Experiment 4). The absence of prior reports of eye movement responses in the anterior STS region is also consistent with this interpretation (Grosbras et al., 2005; Bakola, Gregoriou, Moschovakis, Raos & Savaki, 2007). By contrast, even minute eye movements elicit responses in early visual cortex (Dimigen, Valsecchi, Sommer & Kliegl, 2009) and an eye movement account would seem to account well for the pattern of ipsilateral univariate selectivity I observed in Experiment 1, with leftward and rightward head turns producing responses in left and right early visual cortex, respectively. Notably, this ipsilateral pattern of effects reverted to the expected contralateral response preference in the univariate analysis of Experiment 4, even though left-right head turn classification in early visual cortex was significant in both experiments. These results suggest that the classification effects in the two data sets were driven by distinct large-scale univariate effects: a primarily eye movement-related response in Experiment 1 and a visually-evoked response in the follow-up Experiment 4.

The pervasive tendency for volunteers to follow social attention cues points to an intriguingly close link between action and perception in this system, which is worthy of further enquiry. Previous investigators found that static gaze cues also evoke small eye movements in the perceived gaze direction (Mansfield et al., 2003). Indeed, two of the five volunteers who were tested with eye tracking in the current study were unable to consistently suppress eye movements in response to the head turns, even in the presence of a fixation cross and strong instructions to maintain fixation. Although interesting in their own right, these eye movement effects also suggest that investigators who seek to isolate effects of perceived gaze direction would be well advised to monitor the volunteer's own gaze.

3.8.5 Effects in MT+

Previous studies have found that socially relevant motion engages MT (Puce et al., 1998; Watanabe, Kakigi, Miki & Puce, 2006). Consistent with this literature, I observed a univariate response preference for heads relative to ellipsoids in bilateral superior temporal regions likely corresponding to MT+. Despite this category preference

for heads relative to ellipsoids, I found no evidence that response patterns in this region distinguish head turn direction. In previous studies that attempted to decode motion directions, direction sensitivity was weaker in MT than in earlier visual areas (Kamitani & Tong, 2006; Seymour et al., 2009), which the authors attribute to MT's smaller anatomical size compared with earlier visual areas. Although electrophysiological data suggest considerable direction sensitivity in both MT and early visual cortex (Snowden, Treue & Andersen, 1992), such response properties may interact with area size when measured with coarse-grained methods such as fMRI, thus producing apparently weaker or non-significant effects in smaller areas (Bartels, Logothetis & Moutoussis, 2008). Note also that both the absence of a functional MT localiser and the use of weaker, more transient motion stimuli may have rendered the analysis less sensitive to direction-sensitive responses in MT+, compared with previous studies (Kamitani & Tong, 2006; Seymour et al., 2009). Thus, I do not wish to exclude the possibility that head turns produce direction-sensitive MT+ responses, although I was unable to find evidence for this.

3.8.6 Conclusions

I have presented evidence that response patterns in human right anterior STS/STG distinguish between leftward and rightward dynamic head turns. Such direction sensitivity was not detected for physically matched ellipsoid control stimuli. The anterior location of this effect is consistent with evidence from macaque electrophysiology (Perrett et al., 1985; Hasselmo et al., 1989) but does not co-localise with regions showing greater univariate responses to heads than to ellipsoids. In this respect, multivariate pattern approaches show great promise in linking evidence from single neurons in the macaque to large-scale responses in human fMRI.

Chapter 4

A head view-invariant representation of gaze direction in anterior superior temporal sulcus

4.1 Introduction

Humans show a remarkable ability to discriminate others' gaze direction, even though a given direction can be conveyed by many physically dissimilar configurations of different eye positions and head views. For example, eye contact can be signalled by a rightward glance in a left turned head or by direct gaze in a front-facing head (Figure 4.1a). Such acute gaze discrimination implies considerable perceptual invariance. Previous human research found that STS responds preferentially to gaze shifts (Nummenmaa & Calder, 2009) but the underlying representation that supports such general responsiveness remains poorly understood.

Previous reports of STS involvement in perception of gaze and head view used faces in which eye position or head view were manipulated in isolation (Experiment 1, Chapter 3, see also Calder et al., 2007; Natu et al., 2010). Such designs cannot address the issue of view-invariant coding of gaze because the degree of eye position or head view change defines the degree of gaze direction change. Here, I used computer-generated faces where 9 different gaze directions were conveyed through multiple, physically-dissimilar configurations of different head views and eye position. This de-

sign makes it possible to disentangle fMRI responses consistent with a representation of view-invariant gaze direction from responses related to the faces' other physical features (Todorović, 2006). Moreover, previous attempts to identify view-invariant gaze codes using conventional univariate analysis of smoothed fMRI data have produced inconsistent results and did not observe gaze effects in STS (Pageler et al., 2003; George et al., 2001). This is perhaps unsurprising since macaque STS neurons that are selective for head view and gaze direction are organised into small patches (Perrett et al., 1984; Wang et al., 1998) beyond the likely resolution of conventional fMRI analysis methods. Recently, MVPA has been used to identify other visual representations thought to be coded at similarly small spatial scales, including direction-specific motion responses in early visual cortex (Kamitani & Tong, 2006; Seymour et al., 2009). Here, I applied novel MVPA methods (representational similarity analysis, Kriegeskorte et al., 2008) to high-resolution fMRI data in order to reveal response pattern codes for view-invariant gaze direction.[1]

4.2 Experiment 5: View-invariant gaze direction responses in human fMRI

4.2.1 Materials and methods

4.2.1.1 Participants

Twenty-three right-handed volunteers with normal or corrected to normal vision were recruited for the study. Volunteers provided informed consent as part of a protocol approved by the Cambridge Psychology Research Ethics Committee. Data from five volunteers were removed from the analyses reported here: two failed to complete the experiment, two fell asleep and displayed excessive head motion and one failed to maintain fixation (Section 4.2.1.6). This left 18 volunteers (5 male, mean age 24, range 18-36).

[1]The research described here was published in *Current Biology* (Carlin, Calder, Kriegeskorte, Nili & Rowe, 2011). It is reprinted here in chapter form (see Appendix B for journal form).

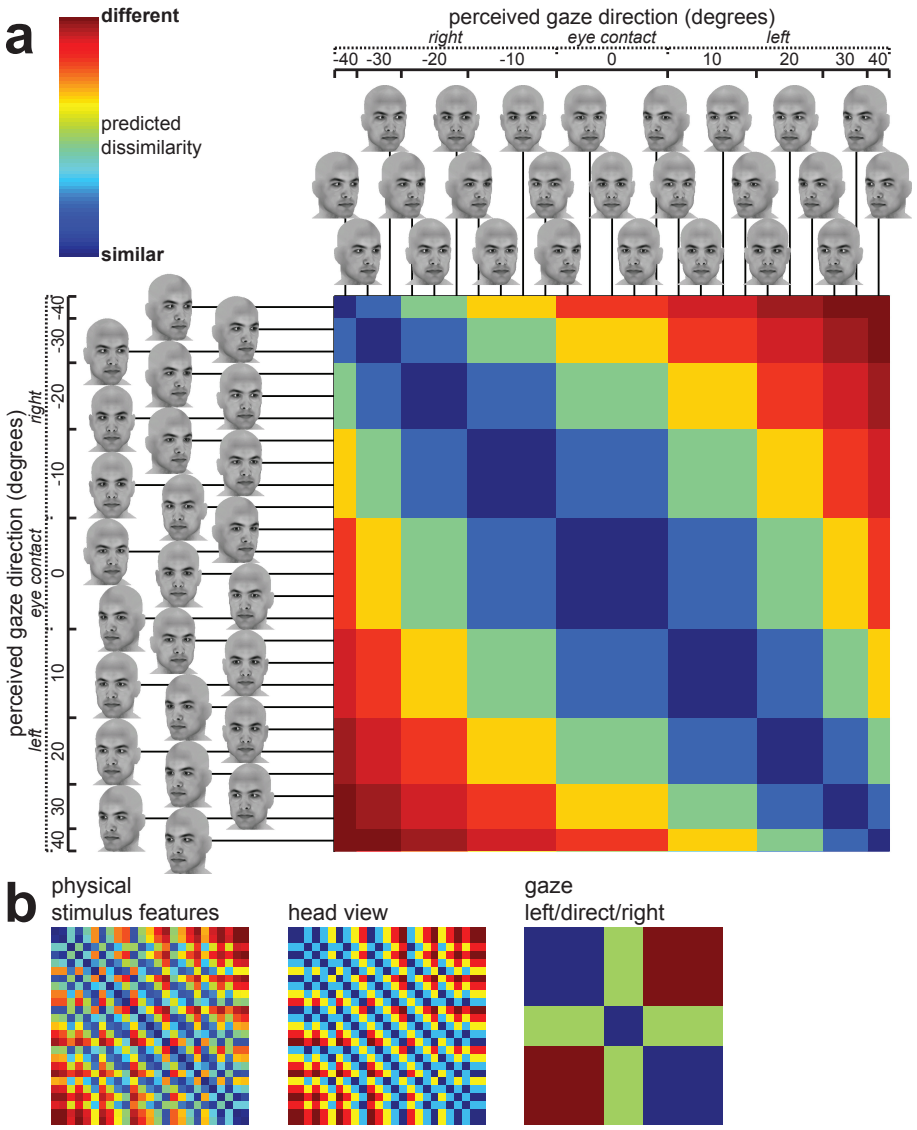


Figure 4.1: Stimuli and predicted RDMs. (**a**) Predicted view-invariant gaze direction RDM across the 25 computer-generated faces. The faces are sorted according to the 9 distinct gaze directions in the stimulus set (left $40°$ to right $40°$ rotation), which were created by incrementally varying head view and eye position relative to the head (5 increments between left $20°$ and right $20°$ for both). (**b**) Predicted RDMs for the same faces based on alternative accounts of the data corresponding to their physical stimulus features ($1 - r$ across image grey scale intensities), head view and qualitative gaze direction (left/direct/right gaze) ignoring quantitative differences between angles of left and right gaze. Dissimilarity matrices in (b) are sorted as in (a).

4.2.1.2 Stimulus design and presentation

I used Poser 6 (Curious Labs Inc. Santa Cruz, CA) to create grey scale face images of two identities, each displaying 25 head-gaze configurations. Each face varied in horizontal head view (5 increments from left 20° to right 20°), horizontal eye position relative to the head (same increments as for head view) and identity (2 faces). The faces were processed in Matlab to achieve similar luminance histograms and were cropped to ensure that each face appeared in a similar retinal area. Cropping was achieved with a smooth border and the resulting face was superimposed on a background texture that varied across conditions and across repetitions of the same face (Figure 4.2). The background textures were created by Fourier-scrambling each of the 50 faces separately. The inclusion of the background texture served to reduce the influence of low-level physical differences between the conditions and to increase the difficulty level of the one-back behavioural task. Stimuli were back-projected onto a screen in the scanner which volunteers viewed via a tilted mirror. The stimuli extended $6 \times 6^{\circ}$ visual angle including the background texture and approximately 3° horizontally by 4° vertically without it. The experiment was controlled using Matlab and the Psychophysics toolbox (Brainard, 1997).

4.2.1.3 Experimental design

Volunteers carried out a one-back matching task while viewing the gazing faces in a rapid event-related fMRI experiment (Figures 4.1, 4.2). Volunteers maintained fixation on a central cross. The faces were presented so that the cross fell on the bridge of the nose of each face to minimise eye movements during the task. The 25 head/eye position configurations were posed by 2 identities (50 images total). Each was presented 3 times in 5 independently randomised sets (150 experimental trials presented over 11 minutes per set; 750 trials in total over 55 minutes). Each trial comprised a face (1 s) followed by an inter-trial interval (2.9 s). Fifteen randomly selected trials in each set were immediately followed by a second presentation of an identical face (75 added trials in total). Volunteers were asked to identify repetitions with a button response before the onset of the next trial (one-back task). Response trials were equally sampled from all head/eye position configurations and were modelled with a separate regressor of no interest in the first level fMRI model. At the end of each set, volunteers viewed a

feedback screen (20 s) that summarised their hit rates and false alarm rates for that set.

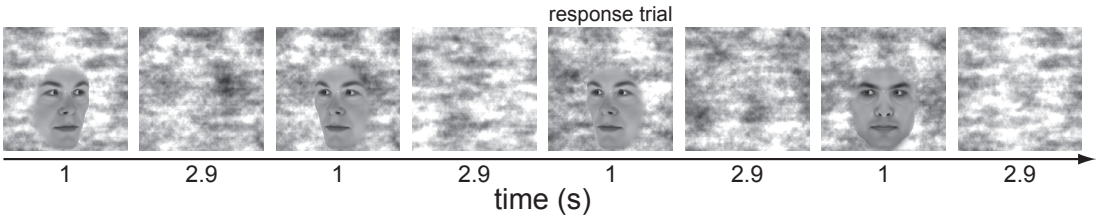


Figure 4.2: An example trial sequence from Experiment 5.

4.2.1.4 Imaging acquisition

Scanning was carried out at the MRC Cognition and Brain Sciences Unit, Cambridge, UK, using a 3 Tesla TIM Trio MRI scanner (Siemens, Germany), with a head coil gradient set. Functional data were collected using high resolution echo planar T2*-weighted imaging (40 oblique axial slices, TR 2490 ms, TE 30 ms, in-plane resolution 2×2 mm, slice thickness 2 mm plus a 25% slice gap, 192×192 mm field of view). The acquisition window was tilted up approximately 30° from the horizontal plane to provide complete coverage of the occipital and temporal lobes. All volumes were collected in a single, continuous run for each volunteer. The initial 6 volumes from the run were discarded to allow for T_1 equilibration effects. T_1-weighted structural images were also acquired (MPRAGE, 1 mm isotropic voxels).

4.2.1.5 Imaging analysis

Pre-processing of the fMRI data was carried out using SPM5. Structural volumes were segmented into grey and white matter partitions and normalised to the MNI template using combined segmentation and normalisation routines. All functional volumes were realigned to the first non-discarded volume, slice timing corrected and co-registered to the T_1 structural volume. The functional volumes remained unsmoothed and in their native space for volunteer-specific generalised linear modelling. Each set was modelled with a separate set of regressors for each head/eye configuration (25, collapsing across the 2 face identities), false alarms, and repeat trials. I also included scan nulling regressors to eliminate the effects of excessively noisy volumes (Lemieux et al., 2007;

Rowe et al., 2008). The experimental predictors were convolved with a canonical haemodynamic response function and contrast images for each individual condition against the implicit baseline were generated based on the fitted responses. The resulting T contrast volumes were grey matter masked using the tissue probability maps generated by the segmentation processing stage, and were used as inputs for representational similarity analysis.

Representational similarity analyses were carried out using custom code developed using Python and PyMVPA (Hanke et al., 2009). The voxels within each searchlight and each set were correlated across conditions ($1 - r$, Pearson) and the resulting 1-correlation matrix was averaged across the 5 sets to produce a final response pattern RDM for that searchlight. The data dissimilarities were then compared to a set of hypothesis-based predictors using the Spearman rank correlation or partial Spearman rank correlation. In all cases, the resulting correlation coefficient was Fisher transformed and mapped back to the central voxel in the searchlight, yielding a descriptive individual subject map that was entered into a group analysis. This two-stage summary statistics procedure resembles that used in conventional univariate fMRI group analysis (Holmes & Friston, 1998). The individual subject maps were normalised to the MNI template and were smoothed to overcome errors in inter-subject alignment (10 mm FWHM). The resulting volumes were entered into a permutation-based random effects analysis (SnPM, 10000 permutations, 10 mm FWHM variance smoothing Nichols & Holmes, 2001). The use of non-parametric tests avoids distributional assumptions regarding the nature of the descriptive maps and avoids inherent problems in applying standard SPM 5 FWE correction based on random Gaussian fields to discontinuous grey matter-masked data.

Based on previous evidence for right-lateralised gaze responses in human STS (Experiment 1, Chapter 3, see also Calder et al., 2007; Pelphrey et al., 2003), I report all p values in the primary analysis corrected for multiple comparisons within the anatomically-defined right STS region ($p < 0.05$, FWE, Figure 4.3, 4598 voxels). For completeness I also carried out exploratory analyses of left STS and the full grey matter-masked volume.

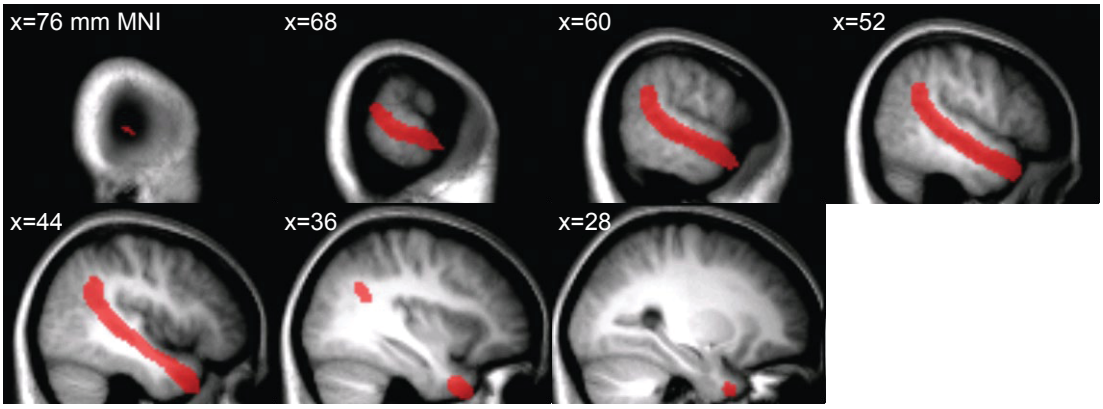


Figure 4.3: Spatial extent of the right STS anatomical mask, shown overlaid on the sample's mean T_1 volume.

4.2.1.6 Eye tracking

Volunteers' eye movements were monitored in the scanner using an infra-red video-based eye tracker (50 Hz acquisition, Sensomotoric Instruments, Germany). Successful calibrations were obtained for 10 volunteers out of the final sample of 18. The remaining volunteers were excluded from the eye tracking analysis. On-line visual inspection of the eye tracking monitor suggested that these volunteers were maintaining their gaze at the fixation cross. Eye tracking data were analysed using custom code developed in Matlab.

4.2.2 Results

4.2.2.1 Eye tracking effects

To measure stimulus-induced eye movements, I analysed how the horizontal and vertical fixation position shifted between the start and the end of each stimulus presentation. This fixation shift was analysed using ANOVAs for each individual volunteer. I used a one-way ANOVA where the faces were labelled according to perceived gaze direction. One volunteer showed an effect of gaze direction on horizontal fixation shifts ($F_{(711,8)} = 2.37$, $p = 0.016$). This volunteer was removed from analysis of the fMRI data. No other horizontal or vertical fixation shift effects were significant in single subjects or in a group analysis. Thus, the volunteers were able to comply with the

instruction to maintain fixation on the central cross.

4.2.2.2 Behavioural performance

Volunteers carried out a one-back face matching task whilst in the scanner. The task instruction was to respond to any repetition of the same face (same identity and head view/eye position configuration) while ignoring the scrambled backgrounds, which did not repeat. Accuracy on the one-back repetition detection task was relatively high across the sample (mean 77%, standard error 3%), with low false alarm rates (mean 4.5% of trials, standard error 2.3%) and high sensitivity (mean d' 2.52, standard error 0.13). The large number of different head view/eye position configurations (25) relative to the number of response trials (75 per volunteer) meant that there was insufficient behavioural data available to model each of the 25 configurations separately. Thus, I pooled the available response trials according to gaze direction and calculated accuracy scores for each of the 9 gaze directions. Repeated-measures ANOVA revealed no significant accuracy effects of gaze direction ($p > 0.23$), suggesting that attentional or performance differences did not confound the fMRI analysis.

4.2.2.3 Right STS gaze codes are invariant to head view and physical stimulus features

Response patterns in right anterior ($p = 0.013$) and posterior ($p = 0.006$) STS showed a consistent relationship with the view-invariant gaze direction predictor (Figure 4.4a). Complementary functional ROI analyses of right STS revealed moderate independently estimated effect sizes in these regions ($r = 0.39$ for anterior STS, $r = 0.42$ for posterior, Figure 4.5). Although these effects suggest that both regions code the direction of another's gaze, it was important to correct for unavoidable correlations between the view-invariant gaze direction predictor and alternative predictors derived from the faces' physical stimulus features ($1 - r$ across image grey scale intensities) or head view (correlation between (i) gaze direction and physical stimulus features $r = 0.37$, (ii) gaze direction and head view $r = 0.36$; Figure 4.1b). Note that the relationship with both is because the faces' physical stimulus features were almost entirely explained by head view ($r = 0.99$).

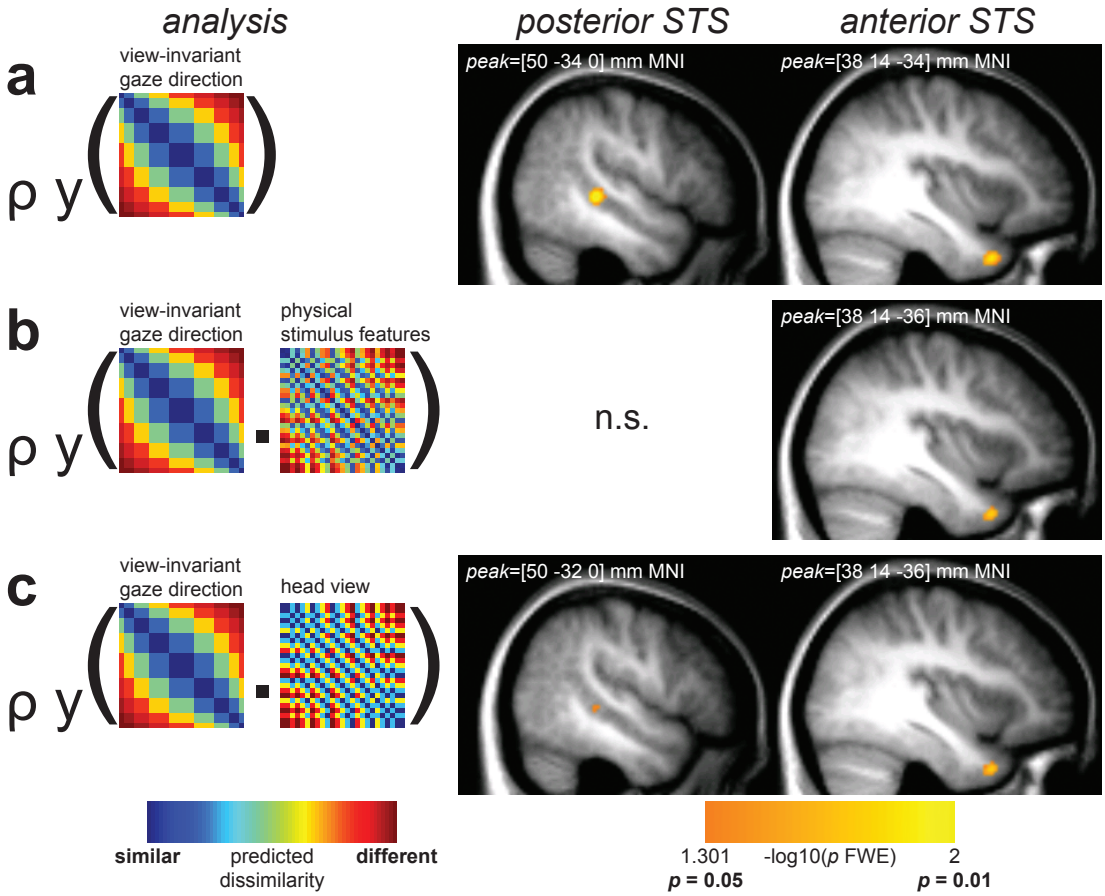


Figure 4.4: Regions with consistent pattern responses (Spearman correlation or partial correlation) across volunteers ($N = 18$, $p < 0.05$, FWE-corrected for right STS, Figure 4.3). Effects are shown overlaid on the sample's mean structural volume. (**a**) Response pattern dissimilarities in anterior and posterior STS are explained by the view-invariant gaze direction predictor. (**b**) Gaze direction responses in anterior STS alone are found for the same predictor when controlling for physical stimulus features. (**c**) Similarly, gaze direction responses in anterior STS for the view-invariant gaze predictor are unaffected when controlling for head view while responses in posterior STS are reduced.

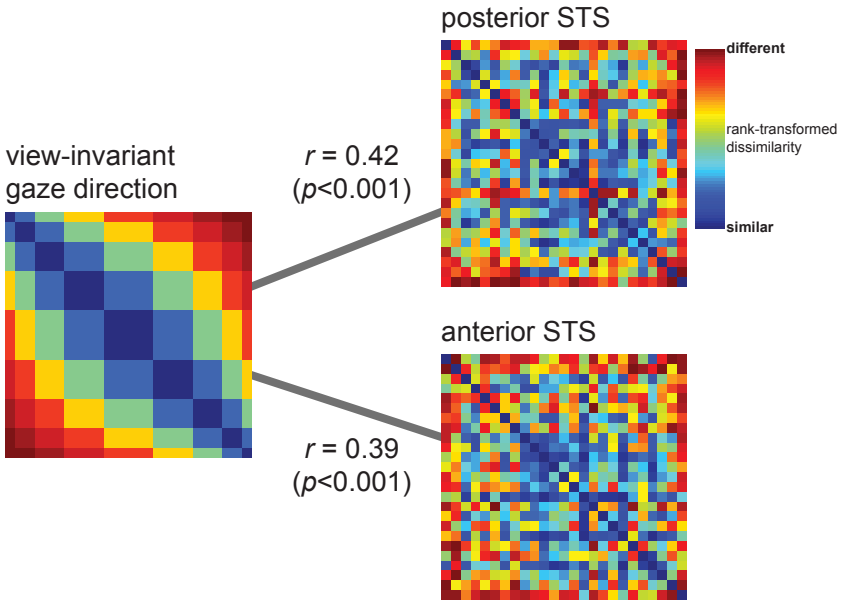


Figure 4.5: Independently estimated correlation between the view-invariant gaze direction predictor and response pattern dissimilarities in anterior and posterior right STS regions. I defined ROIs using a leave-one-set-out procedure. I carried out a group analysis (similar parameters as main analysis) separately for the ROI-defining data in each unique split (4 of 5 sets) of the data to identify response pattern dissimilarities that were explained by view-invariant gaze direction. Responses to each set were estimated in five separate first-level models with 7 discarded volumes (17.43 s) separating each model to ensure independent estimates. Statistical thresholds for ROI definition varied between splits ($p < 0.01$ to $p < 0.05$, uncorrected). The only regions that appeared consistently across splits were anterior STS (mean $[33.2, 10.0, -41.2]$ mm MNI, standard deviation $[1.0, 5.1, 1.0]$) and posterior STS (mean $[46.4, -29.8, 4.0]$ mm MNI, standard deviation $[3.2, 3.5, 4.2]$). To better accommodate alignment errors across volunteers, I identified the volunteer-specific peak within a 10 mm radius of each group peak using ROI-defining data only. Subsequent tests of the identified ROIs were carried out separately for each split (e.g., ROIs defined using sets 1-4 were tested using set 5). I generated the illustrated response pattern dissimilarities for anterior and posterior STS by first averaging each volunteer's dissimilarities for each ROI across the 5 independent test splits and then averaging the resulting ROI dissimilarity matrices across volunteers. It can be seen that both anterior and posterior STS showed consistent effects of view-invariant gaze direction in the independent test data (p values were defined using a permutation test where the order of the matrices were shuffled without replacement 10000 times, Kriegeskorte et al., 2008).

To exclude the contribution of these additional facial properties I computed a further correlation between the view-invariant gaze direction predictor and the response pattern dissimilarities, this time partialling out any correlation between physical stimulus features and the response pattern dissimilarities (partial Spearman correlation). Only the perceived gaze direction effect in anterior STS remained significant when the influence of physical stimulus features was removed ($p = 0.018$, Figure 4.4b). Similarly, removing the influence of head view did not disrupt the effect of the view-invariant gaze direction predictor in anterior STS ($p = 0.016$) but produced only a weakly significant effect in posterior STS ($p = 0.045$, Figure 4.4c). Indeed, posterior STS showed a significant relationship with the predictor derived from the faces' physical stimulus features ($p = 0.048$) and a near-significant relationship with the predictor derived from head view ($p = 0.08$). Thus, gaze direction responses in posterior STS were influenced by physical stimulus features, which corresponded largely to variation in head view while gaze direction responses in anterior STS were invariant to these facial properties.

4.2.2.4 Right STS gaze codes are fine-grained

If gaze codes in STS play a role in supporting perceptual performance, such codes should mirror human sensitivity to fine-grained gaze direction distinctions (Symons et al., 2004). I tested this by comparing the original view-invariant gaze predictor representing 9 gaze directions to a left/direct/right gaze predictor that distinguished between three qualitative gaze directions while ignoring continuous information about the degree to which gaze is averted left or right (Figure 4.1b). Partial correlation analysis showed that the effects of the original view-invariant gaze predictor remained after removing the influence of the left/direct/right gaze predictor (anterior STS: $p = 0.016$, posterior STS: $p = 0.018$, Figure 4.6). Thus, the reported view-invariant gaze direction effects cannot be explained by simpler gaze representations. Instead, gaze direction codes in STS contained fine-grained information about both the direction and the degree to which gaze is averted.

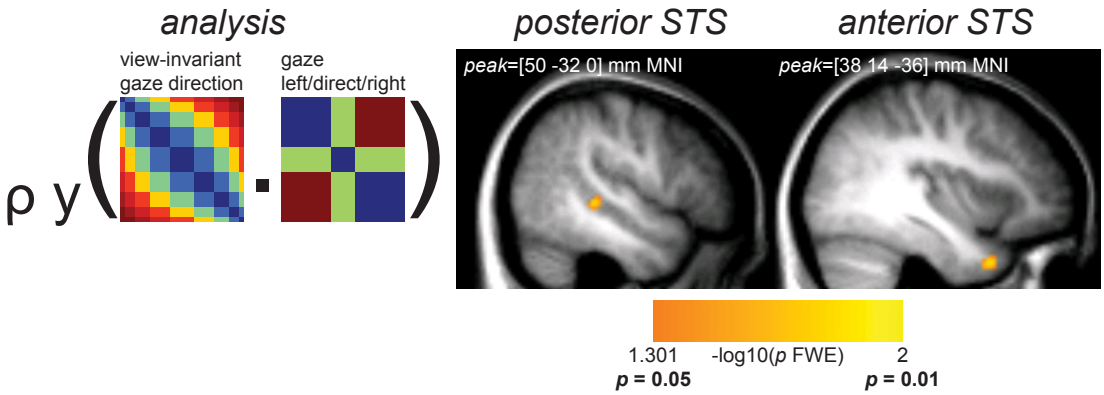


Figure 4.6: Fine-grained gaze direction codes in right STS. Regions with consistent pattern responses (partial Spearman correlation) across volunteers ($N = 18$, $p < 0.05$ FWE). View-invariant gaze direction responses in anterior and posterior right STS remain when the influence of a qualitative distinction between gaze left/direct/right is removed.

4.2.2.5 Gaze codes in left STS and precuneus

An exploratory analysis of left STS revealed similar evidence of view-invariant coding of gaze direction in left anterior STS (Table 4.1). There were no significant effects in left posterior STS ($p > 0.19$). View-invariant representations of gaze direction in anterior STS may therefore be bilateral.

A further analysis of the full grey-matter masked volume also revealed view-invariant gaze direction codes in precuneus, which survived all control analyses reported above (Table 4.1). Precuneus and STS are monosynaptically connected in macaques (Seltzer & Pandya, 1994) and precuneus has previously been implicated in head/gaze following (Laube, Kamphuis, Dicke & Thier, 2011) and in attentional orienting (Cavanna & Trimble, 2006), which suggests that gaze codes here may reflect gaze-cued shifts in attention (Friesen & Kingstone, 1998). Eye tracking analyses suggested that volunteers were fixating well (Section 4.2.1.6), so these precuneus effects are likely driven by covert attentional shifts rather than overt eye movements.

Table 4.1: Analyses of additional regions and tests of the performance of a volunteer-specific gaze discrimination predictor in place of the standard view-invariant gaze direction predictor. Peak MNI coordinates are shown with p values FWE-corrected for regions as indicated by the analysis column.

Analysis	Comparison	Region	p (FWE)	Peak (mm MNI) x	y	z
right STS, 4598 voxels ($p < 0.05$)	View-invariant gaze direction	posterior STS	0.01	50	-34	0
		anterior STS	0.01	38	14	-34
	...partialling out physical stimulus features	anterior STS	0.02	38	14	-36
	...partialling out head view	posterior STS	0.05	50	-32	0
		anterior STS	0.02	38	14	-36
	...partialling out gaze left/direct/right	posterior STS	0.02	50	-32	0
		anterior STS	0.02	38	14	-36
	Physical stimulus features	posterior STS	0.05	50	-32	2
	Volunteer-specific gaze discrimination	posterior STS	0.01	50	-34	0
		anterior STS	0.04	38	12	-36
	...partialling out physical stimulus features	anterior STS	0.05	38	12	-36
		posterior STS	0.06	50	-34	0
	...partialling out head view	posterior STS	0.04	50	-34	0
		anterior STS	0.05	38	12	-36
	...partialling out gaze left/direct/right	posterior STS	0.01	50	-34	0
		anterior STS	0.04	38	12	-36
left STS, 4442 voxels ($p < 0.1$ to illustrate marginally significant effects mirroring those observed in right STS)	View-invariant gaze direction	anterior STS	0.04	-60	-8	-16
	...partialling out physical stimulus features	anterior STS	0.05	-62	-6	-16
	...partialling out head view	anterior STS	0.05	-62	-6	-16
	...partialling out gaze left/direct/right	anterior STS	0.03	-60	-8	-16
		middle STS	0.07	-66	-26	-8
	Volunteer-specific gaze discrimination	anterior STS	0.05	-60	-6	-16
	...partialling out physical stimulus features	anterior STS	0.07	-62	-6	-14
	...partialling out head view	anterior STS	0.05	-62	-6	-14
	...partialling out gaze left/direct/right	anterior STS	0.03	-60	-8	-16
		middle STS	0.06	-66	-26	-8
whole brain analysis, 134174 voxels ($p < 0.05$)	View-invariant gaze direction	precuneus	0.01	4	-58	30
		cerebellum	0.04	-16	-76	-50
		cingulate gyrus	0.05	2	-22	32
	...partialling out physical stimulus features	precuneus	0.02	4	-58	30
	...partialling out head view	precuneus	0.01	4	-58	30
		cerebellum	0.05	-18	-76	-50
	...partialling out gaze left/direct/right	precuneus	0.03	4	-58	30
		cerebellum	0.02	-16	-76	-50
		occipital pole	0.04	-10	-98	2
	Volunteer-specific gaze discrimination	precuneus	0.01	4	-58	30
		posterior STS	0.04	50	-34	0
	...partialling out physical stimulus features	precuneus	0.03	4	-58	30
	...partialling out head view	precuneus	0.02	4	-58	30
		cerebellum	0.04	-18	-76	-50
	...partialling out gaze left/direct/right	precuneus	0.01	6	-58	30
		posterior STS	0.03	50	-34	0
		anterior STS	0.04	-60	-12	-26

4.3 Experiment 6: View-invariant gaze discrimination in behavioural performance

My experimental design assumes that perceived gaze direction can be approximated by the sum angle of head view and eye position relative to the head (Figure 4.1, Todorović, 2006). However, human gaze discrimination performance can be subtly biased by head view (Gibson & Pick, 1963; Gamer & Hecht, 2007). I therefore carried out a follow-up behavioural experiment to assess whether the standard view-invariant gaze predictor I used was a good match for the volunteers' individual gaze discrimination performance.

4.3.1 Materials and methods

At the end of the fMRI study (Experiment 5), all volunteers carried out a follow-up gaze direction discrimination experiment outside the scanner. Volunteers viewed the stimuli from Experiment 5 from a fixed position in a head rest. On each trial, volunteers were asked to indicate the perceived direction of gaze by rotating a pointer in the horizontal plane; this pointer was positioned in front of the screen directly underneath the stimulus. Separate RDMs were then generated from each volunteer's discrimination data by computing the difference between the mean pointer positions for each pair of head view/eye position configurations. The relationship between each volunteer's perceptual discrimination and the predicted dissimilarity structures was then estimated using a similar procedure to the main fMRI analysis in Experiment 5. Inference was carried out at the group level, using confidence intervals based on bootstrap testing of the median (*bootci* function in Matlab, 95% interval, 10000 samples).

4.3.2 Results

Volunteers' gaze discrimination performance was well explained by the standard view-invariant gaze direction predictor (median Spearman $r = 0.90$, 95% confidence=0.87-0.93, bootstrap test) and this relationship survived removing the influence of each of the alternative predictors discussed above (Figure 4.7). The standard gaze direction predictor was thus a good approximation for volunteers' perceived gaze direction in this stimulus set.

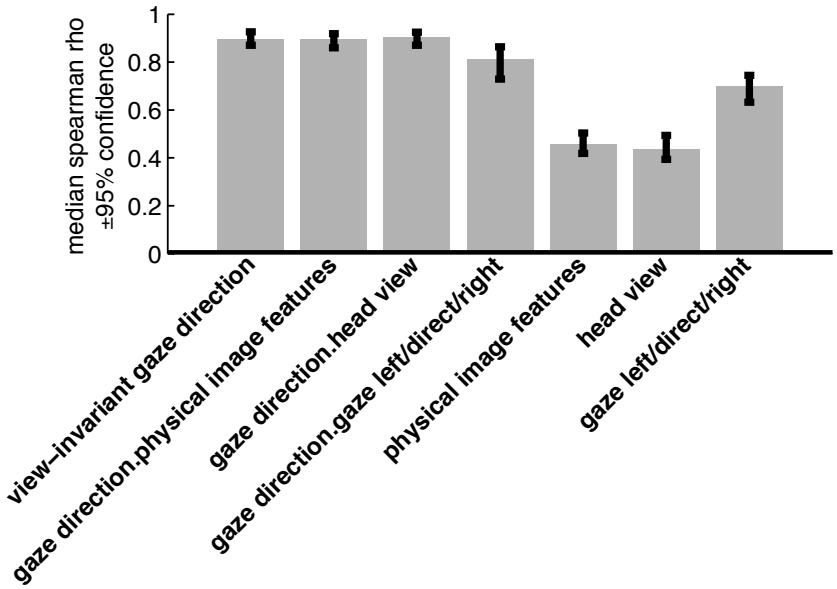


Figure 4.7: Behavioural gaze direction discrimination. Median Spearman correlations (bars 1,5-7) and median partial Spearman correlations (bars 2-4) across the volunteers ($\pm 95\%$ bootstrap confidence intervals). The volunteers' gaze discrimination performance was most strongly correlated with the view-invariant gaze direction predictor. Although performance was also moderately correlated with physical image features and head view, the strength of the relationship between discrimination performance and the view-invariant gaze direction predictor is relatively unaffected by partialling out the influence of these alternative predictors.

I also repeated the fMRI analyses in Experiment 5 using the volunteer-specific gaze discrimination predictors in place of the standard view-invariant gaze direction predictor and obtained comparable results (Table 4.1). Thus, volunteers' gaze discrimination performance was well approximated by the standard view-invariant gaze predictor and the neural responses to the gazing faces were similarly explained by the standard and volunteer-specific gaze predictors.

4.4 General discussion

Experiment 5 provides the first evidence that human anterior STS contains a fine-grained, view-invariant code of perceived gaze direction. I also observed similar gaze effects in precuneus, which may reflect attentional orienting responses to gaze (Friesen & Kingstone, 1998). These results do not rule out the existence of view-specific codes for particular head-gaze configurations, but rather demonstrate that gaze perception is not achieved using such view-specific representations alone. The results are consistent with the hypothesis that gaze perception is achieved through a high-level, view-invariant code of the direction of another's social attention in anterior STS.

The representational content of right posterior STS is distinct from anterior STS. Although the view-invariant gaze predictor also identified this region, this was largely accounted for by the modest correlation between this predictor and the faces' physical facial properties or head view, which showed significant or borderline relationships with the right posterior STS. This is consistent with recent work showing that response patterns in posterior STS can be used to distinguish head view (Natu et al., 2010). The preferential involvement of anterior STS in view-invariant representations of gaze direction was further underlined by the analysis of left STS, which identified the anterior region alone. The results are thus consistent with previous reports that right posterior STS is responsive to different gaze directions and head views (Nummenmaa & Calder, 2009; Natu et al., 2010) but view-invariant gaze direction codes appear most prevalent in anterior STS.

Collectively, the results in this chapter suggest a hierarchical processing stream for gaze perception, with increasing invariance to gaze-irrelevant features from posterior to anterior STS. Such a processing hierarchy would be consistent with recent evidence from neurons responsive to face identity in the macaque temporal lobe (Freiwald &

Tsao, 2010), where invariance to head view increases from middle STS to anterior inferotemporal cortex. Similarly, neurons tuned to specific head views in anterior STS also frequently respond to gaze direction (Perrett et al., 1992, 1985; De Souza et al., 2005) while neurons with head view tunings in middle STS generally do not (De Souza et al., 2005). Such hierarchical progressions toward view invariance may therefore be a general property of high-level face representations, whether these hierarchies serve to extract face identity or the direction of another's gaze.

In conclusion, response patterns in human anterior STS are not coded according to any readily observable visual face features but rather according to the direction of another person's gaze, irrespective of head view.

Chapter 5

Functional MRI responses to gaze direction and head view in macaque superior temporal sulcus

5.1 Introduction

The human fMRI experiments in the previous chapters were motivated by evidence that single cells in macaque STS code the direction of gaze, head view and head turn (Hasselmo et al., 1989; Perrett et al., 1992, 1985). The process of generating hypotheses for human fMRI experiments based on evidence from macaque electrophysiology involves making two assumptions. First, that neuronal spiking translates into fMRI signals; this correlation is known to be imperfect (see e.g., Sirotin & Das, 2009; Logothetis, 2008). Second, that similar anatomical regions carry out similar functions in humans and macaques; such homology is also imperfect. For instance, motion area MT is found in macaque posterior STS while the homologous area in humans is found in posterior inferior temporal sulcus (Huk, Dougherty & Heeger, 2002; Orban et al., 2004; Nelissen, Vanduffel & Orban, 2006).

In this chapter, I describe fMRI experiments that investigated whether fMRI responses in macaque STS discriminate the direction of gaze and head view. By using similar fMRI and MVPA methods as the previous human experiments, these experiments allow closer comparisons between gaze representations in humans and macaques.

There is evidence that single cells in macaque STS are sensitive to perceived gaze direction (Perrett et al., 1985; De Souza et al., 2005), but only one previous macaque study has tested if responses to gaze direction are also present in the haemodynamic signals that fMRI measures. Hoffman et al. (2007) found increased amygdala responses to averted relative to direct gaze, but no direction-sensitive effects in STS. Although these findings are inconsistent with homologous gaze direction representations in humans and macaques, this study is subject to two limitations. First, comparisons between direct and averted gaze may not distinguish direction-specific representations of gaze from the attentional capture mechanisms of eye contact (see also Section 1.5.2, Emery, 2000; Shepherd, 2010). Second, this study used conventional univariate comparisons between gaze directions. Previous human fMRI studies suggest that fMRI adaptation or MVPA techniques may be necessary to reveal anterior STS involvement in gaze direction coding (Experiment 1, Chapter 3, Calder et al., 2007). The principal aim of Experiment 7 was to test whether classifier MVPA could be used to distinguish gaze direction based on macaque fMRI responses. Furthermore, given the known role of macaque STS in perception of motion (Orban et al., 2004; Nelissen et al., 2006; Mikami, Newsome & Wurtz, 1986a) and faces (Tsao et al., 2003; Pinsk et al., 2005; Tsao et al., 2006), I sought to explore how motion and face sensitivity relates to gaze direction sensitivity.

Although these results are preliminary, they provide initial evidence that gaze direction and head view can be distinguished based on STS response patterns using macaque fMRI.[1]

[1]The experiments in this chapter were conducted in collaboration with Doris Tsao and Ralph Adolphs, California Institute of Technology, using the following division of labour. The design of the experiments was decided collectively. Doris Tsao acquired macaque fMRI data and carried out initial pre-processing. I designed the stimuli, created the experiment presentation script and carried out all reported univariate modelling and MVPA of the data.

5.2 Experiment 7: Gaze direction responses in macaque STS

5.2.1 Methods

5.2.1.1 Participants

Two adult (age 4 years) male rhesus macaques (*macaca mulatta*) were used in the experiment. Both subjects had been scanned in many previous fMRI experiments. Neither subject had received any previous training in behavioural tasks involving discrimination of faces or gaze direction. All methods were consistent with US National Institutes of Health Guidelines, including the US National Institutes of Health Guide for Care and Use of Laboratory Animals.

5.2.1.2 Experimental design

I used a block design (24 s on/24 s off). The blocks were longer than the optimal block duration in human BOLD fMRI (16 s on/16s off, Henson, 2003) because the subjects were scanned using a monocrystalline iron oxide nanoparticle (MION) contrast agent that produces haemodynamic responses with a slower time course than BOLD (Leite, Tsao, Vanduffel, Fize, Sasaki, Wald, Dale, Kwong, Orban, Rosen, Tootell & Mandeville, 2002). The sluggish nature of the MION response also led to the decision to use a block rather than rapid event-related design. The experiment included blocks for 3 gaze direction conditions (left, direct, right) and additional blocks used to localise motion- and face-sensitive areas.

The subjects viewed random sequences of the 24 s experimental blocks (even blocks) alternating with 24 s baseline blocks of scrambled object images (odd blocks). A fixation point was superimposed on each stimulus. The subjects had previously been trained to maintain fixation for juice reward. Eye tracking data were acquired using an infra-red system and juice reward was administered automatically contingent on fixation performance (Tsao et al., 2003).

Repeating a single gaze direction for an entire block would have produced a highly artificial percept so each block contained alternating presentations of a direction of interest (left, direct or right) and a range of alternative gaze directions (not including

the other directions of interest) at a rate of 2 Hz (48 stimulus presentations per block). The paradigm is illustrated in Figure 5.1.

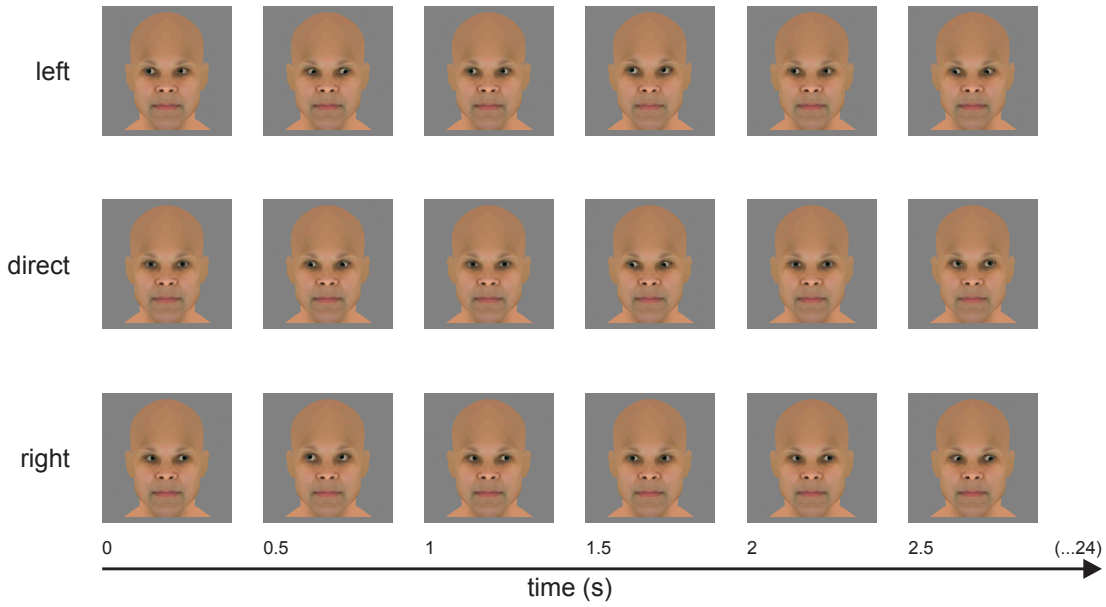


Figure 5.1: Example blocks for the three gaze conditions in Experiment 7. In each condition the gaze direction alternates between the direction of interest (odd presentations) and randomly sampled alternative directions (even presentations). The alternative directions were the same across conditions to ensure that differences could be attributed to the directions of interest. When presented in sequence, the percept is of the stimulus person shifting his gaze to multiple locations at an even rate.

The experiment also included separate blocks with expanding and static dots, blocks with direct-gaze faces that varied in identity, blocks with intermixed presentations of other categories (body parts, fruit, tools, hands) and a block where a grey rectangular occluder that covered the eye region moved in the same sequence of displacements as the iris had in the direct gaze block. I used the comparison between moving and static dots to identify motion sensitive areas.

The comparison between faces and other object categories (pooled into one joint non-face object condition) was intended to identify category sensitive areas. However, the face blocks contained highly similar colour images while the non-face blocks contained highly variable grey scale images. This inadequate matching of physical variance between the face and non-face blocks made it impossible to use face/non-

face comparisons to identify face-selective regions since previous studies have found that physical variance strongly influences neural responses (see e.g., Thierry, Martin, Downing & Pegna, 2007; Davidenko, Remus & Grill-Spector, 2011). Similarly, the occluder block was intended to serve as a low-level control for physical displacements caused by gaze shifts, but this block also failed to achieve adequate physical matching with the gaze blocks and it is not discussed further.

Each scanner run comprised two presentations of each gaze block and one presentation each of the other block types (12 experimental blocks and 13 scrambled blocks, 10 min total run duration). Each subject completed 44 runs in total (7 hours 20 minutes total scan time) over 3 consecutive days.

5.2.1.3 Imaging acquisition

Scanning was carried out at the California Institute of Technology (Pasadena, USA) using a 3-T TIM Trio MRI scanner with an AC88 head gradient insert and an 8 channel surface coil. Prior to each scan session, subjects received a MION contrast agent injection into the femoral artery (concentration: 21 mg Fe/ml in saline, dosage: 8 mg Fe/kg). The subjects were seated in a custom-designed primate chair in the sphinx position (Tsao et al., 2003). Functional EPI data were then acquired using a gradient-echo sequence (TR 2 s, TE 30 ms, 1 mm isotropic voxels, 54 axial slices, 96×96 mm field of view, 300 volumes per run). This acquisition provided whole-brain coverage in both subjects, although there was significant drop-out on the ventral surface of the temporal lobes due to susceptibility artefacts. A fieldmap was also acquired to correct for B0-field inhomogeneities (Cusack, Brett & Osswald, 2003).

5.2.1.4 Preprocessing and modelling

Imaging data were pre-processed using a combination of custom-developed software from the Tsao lab and FMRIB software library (FSL) routines (Smith, Jenkinson, Woolrich, Beckmann, Behrens, Johansen-Berg, Bannister, De Luca, Drobnjak, Flitney, Niazy, Saunders, Vickers, Zhang, De Stefano, Brady & Matthews, 2004). Volumes were undistorted (Cusack et al., 2003) and motion corrected (Jenkinson, Bannister, Brady & Smith, 2002).

All analyses were carried out in single subjects. I fit GLMs using SPM 8 (128 s

high pass filter). Models for MVPA were fitted to unsmoothed volumes using a block-specific modelling approach (Mumford et al., 2011). Each of these block-specific models included one regressor for the modelled block, a second regressor that collapsed all other blocks and a session mean (3 regressors in total). The block-specific models were restricted to volumes from the same run as the modelled block to avoid non-independence in subsequently computed T statistics against baseline across runs (caused by shared variance estimates across runs). For further discussion of this method, see Section 2.2.1.1. These T statistic volumes formed the training examples for MVPA.

All model fitting was carried out through convolution with a gamma function that accommodates the temporal characteristics of the MION haemodynamic response (Formula 5.1, $\tau = 8, \alpha = 0.3$, Leite et al., 2002). All contrast vectors were inverted since the MION responsive is negative while the gamma function below is positive.

$$f(t) = \left(\frac{t - t_0}{\tau}\right)^{\alpha} e^{-\frac{t-t_0}{\tau}} \tag{5.1}$$

5.2.1.5 Multivariate pattern analysis

I used linear SVM classifiers for all MVPA, with similar parameters to Experiment 1 (Chapter 3). Training examples were averaged across the first and second halves of the dataset (Section 2.2.1.2) and were Z-scored so that each voxel within a set had a mean of 0 and a standard deviation of 1. All MVPA was carried out inside a spherical searchlight (3 mm radius, Section 2.2.3, Kriegeskorte et al., 2006) that moved through the entire volume. Unlike previous human experiments, training examples were not grey matter masked since no anatomical T_1 volumes were acquired for segmentation (this would have required anaesthesia). Furthermore, I Winsorised all training examples (cap of 2 standard deviations from the mean, Section 2.2.1.2). For inference, searchlight accuracy maps were converted to p maps by means of a binomial test against chance performance. The absence of a T_1 volume precluded unambiguous anatomical definition of the STS region for region of interest definition so instead I used whole-brain analyses at a reduced statistical threshold ($p < 0.001$, uncorrected). Although this relatively liberal threshold does not control the FWE rate at $p < 0.05$, I focus here on replications across hemispheres and across subjects.

5.2.2 Results

5.2.2.1 Classification effects in macaque lateral temporal lobe

I carried out a searchlight classifier analysis to identify regions where fMRI response patterns could be used to distinguish between the three gaze direction blocks (left, direct, right). Two regions in middle and anterior lower bank STS showed above-chance gaze direction discrimination in left and right STS and in both subjects (Figure 5.2).

These results indicate that regions in middle and anterior macaque STS distinguish the direction of gaze. Previous fMRI studies have shown that macaque STS also contains regions that are sensitive to motion (Orban et al., 2004; Nelissen et al., 2006). Consistent with this, additional searchlight classifier analyses showed extensive discrimination between moving and static dots in macaque STS (Figure 5.2). Gaze and motion discrimination overlapped partially in both subjects but this overlap did not appear in any consistent regions across subjects and hemispheres.

5.2.2.2 Classification effects elsewhere in the brain

Most macaque electrophysiology studies of gaze responses have focused on the role of STS cells. The use of macaque fMRI allows a less constrained search for gaze direction effects across the volume. Contrary to previous reports (Hoffman et al., 2007), there was no evidence of gaze discrimination in amygdala. Instead, extensive gaze discrimination was present in posterior regions likely corresponding to early visual areas V1, V2, V3 and V4 (Figure 5.3). Much like in STS, these effects overlapped partially with motion sensitivity.

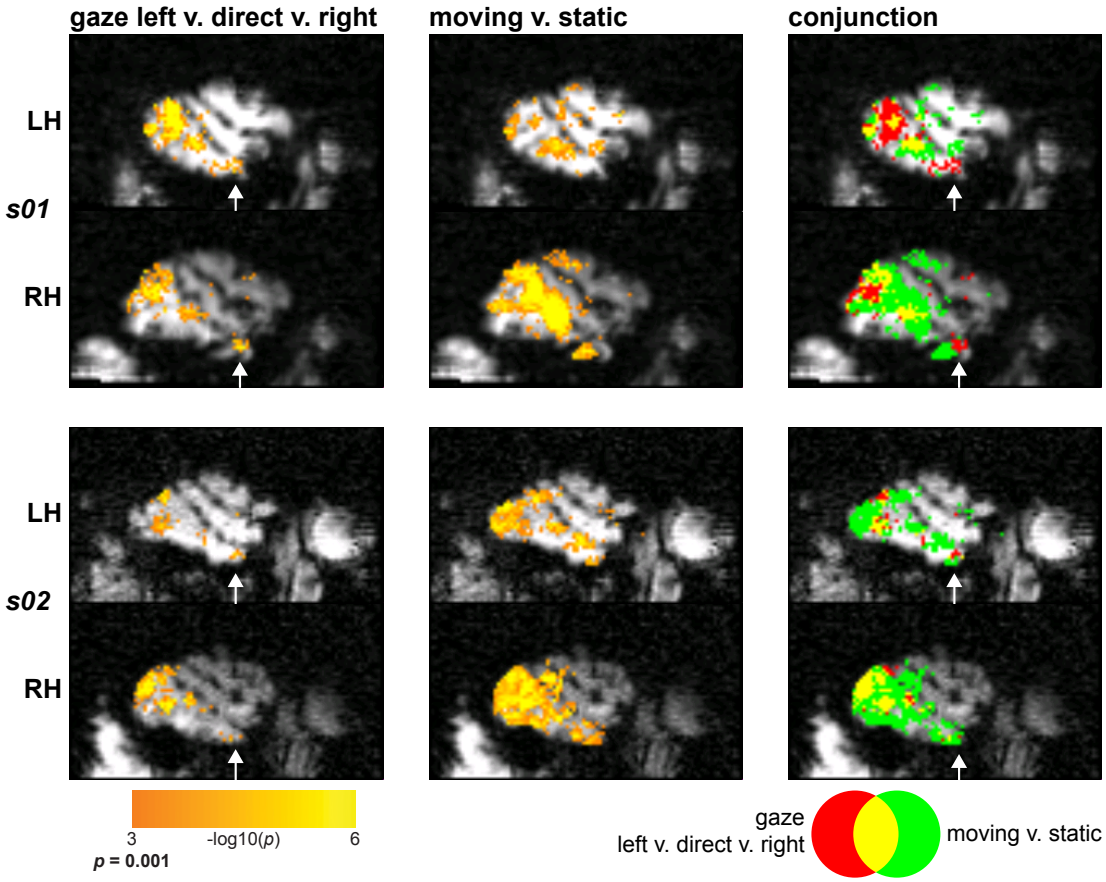


Figure 5.2: Searchlight maps of macaque lateral temporal lobe. Anterior STS effects are illustrated with white arrows. Left panels: Two patches in middle and anterior STS show gaze direction (left v. direct v. right) discrimination. Middle panels: Extensive posterior and middle regions show sensitivity to motion (moving v. static dots). Right panel: A conjunction of the results in left and middle panel shows partial overlap between gaze discrimination and motion sensitivity. The extensive effects in lunate sulcus (posterior to STS) are illustrated more clearly elsewhere (Figure 5.3). The maps are thresholded at $p < 0.001$ (uncorrected).

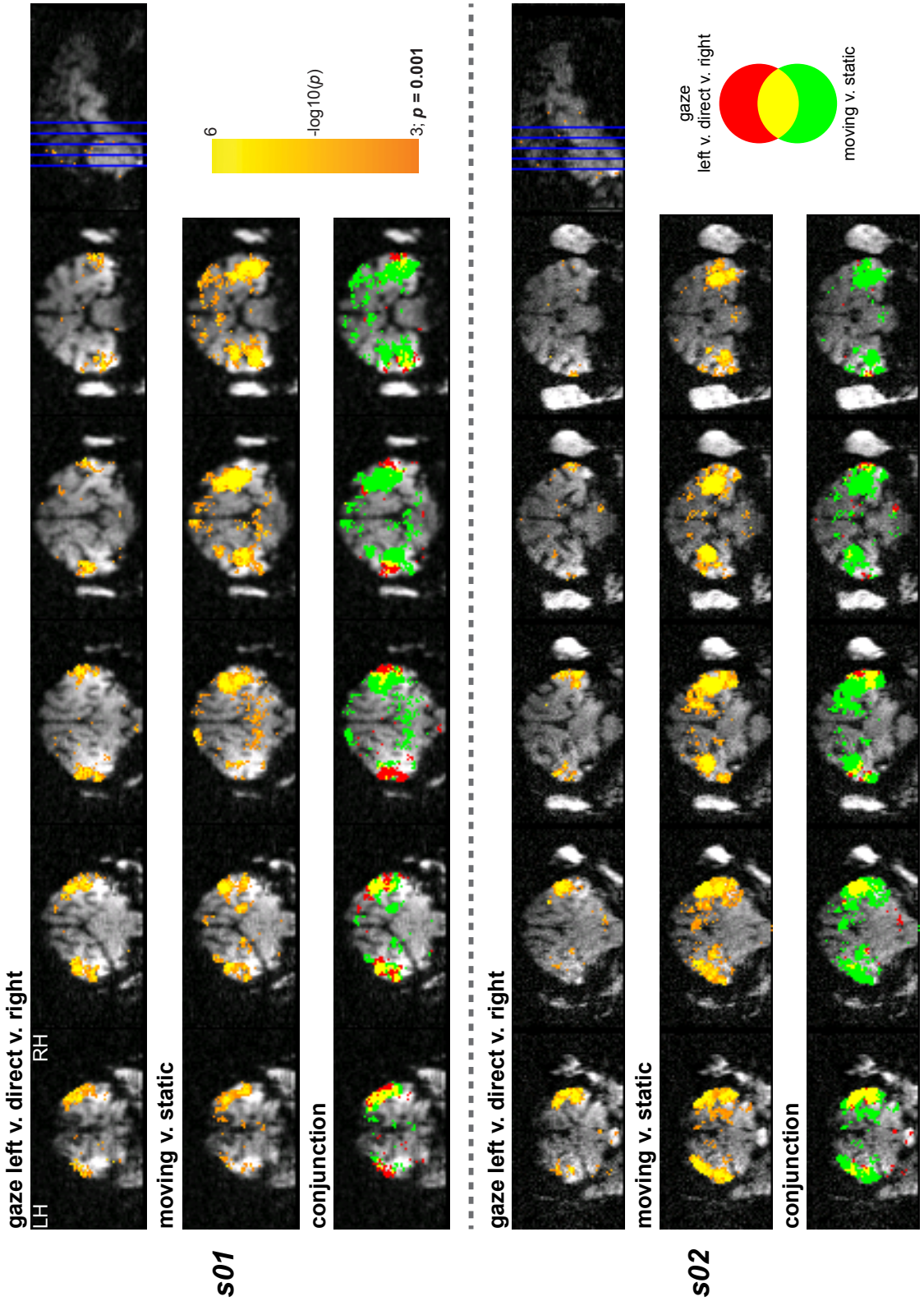


Figure 5.3: Searchlight maps of macaque visual cortex. Both subjects (top and bottom panels) show gaze discrimination effects in regions likely corresponding to early visual areas in posterior temporal and occipital cortex. These effects overlap with motion sensitivity, although motion effects also extend more medially. The maps are thresholded at $p < 0.001$ (uncorrected).

5.2.3 Interim discussion

Experiment 7 showed that macaque anterior and middle STS contains regions where the direction of gaze can be discriminated based on fMRI response patterns. These effects overlapped with motion sensitive regions, which suggests that responses to motion and gaze are either overlapping or organised into distinct sub-regions at a scale beyond the resolution of the searchlight method. Although consistent in both hemispheres of both subjects, the magnitude of the gaze direction effects was moderate. Furthermore, inadequate physical matching between face and non-face blocks meant that this experiment could not be used to test if face sensitivity (whether measured by classifier discrimination against other categories or by univariate response preferences) coincided with gaze direction sensitivity (cf. Experiment 1, Chapter 3).

5.3 Experiment 8: Head view responses in macaque STS

Experiment 8 was designed to improve on Experiment 7 in two ways. First, this experiment used head view changes in place of gaze shifts. Previous investigators have argued that head view is a more salient cue for non-human primates than eye gaze (Kobayashi & Kohshima, 1997; Tomasello, Hare, Lehmann & Call, 2007) so I reasoned that this may translate into correspondingly stronger neural distinctions in macaque STS. Second, I used standard face localiser blocks from the Tsao lab to localise face sensitive areas. The stimuli in these blocks were better matched to the non-face object blocks than those used in Experiment 7.

5.3.1 Methods

The methods in Experiment 8 were similar to Experiment 7. Here I note only features that changed between the two experiments.

5.3.1.1 Participants and experimental design

One macaque participated in Experiment 8 (s01 in Experiment 7). The gaze conditions in Experiment 7 were replaced by comparable head view conditions (Figure 5.4). The

face block in Experiment 7 was replaced with a set of standard grey scale face images from previous localisers in the Tsao lab (Tsao et al., 2003). The experiment comprised 12 runs in total (2 hours total scan time), which were acquired over 2 consecutive days.

The occluder condition in Experimenter 7 was replaced by a head translation condition where a front-facing head shifted in screen location along the same displacements as in the direct head view condition.

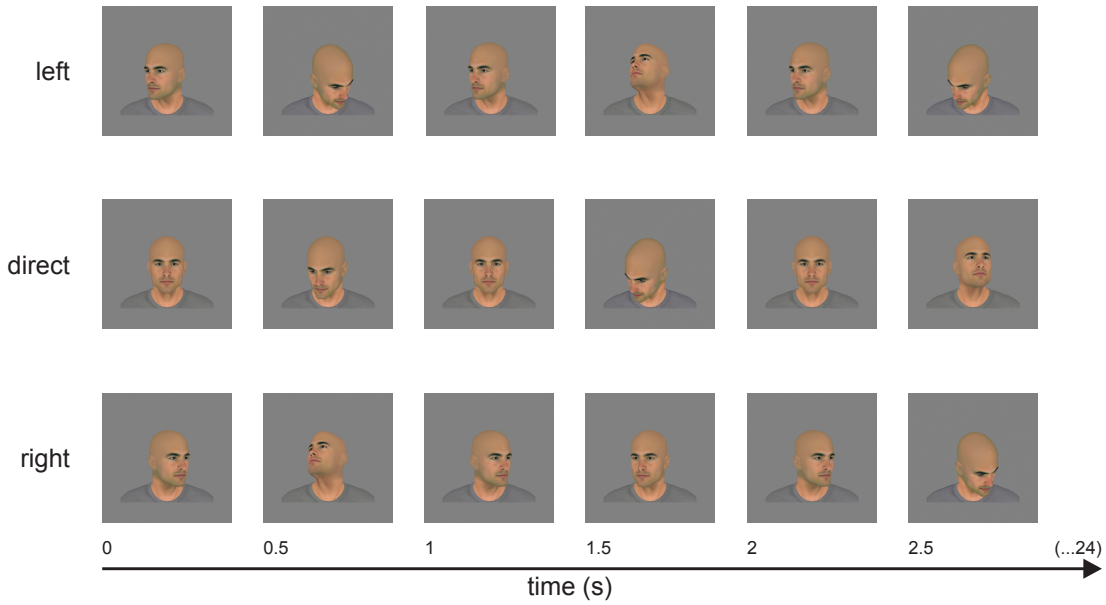


Figure 5.4: Example blocks for the three head view conditions in Experiment 8. In each condition the head view alternates between the view of interest (odd presentations) and randomly sampled alternative views (even presentations). The alternative views were the same across conditions to ensure that differences could be attributed to the directions of interest. When presented in sequence, the percept is of the stimulus person turning his head to multiple locations at an even rate.

5.3.1.2 Imaging acquisition and modelling

Experiment 8 used a slightly different EPI sequence from Experiment 7 (TR 3 s, TE 30 ms, 1 mm isotropic voxels, 54 axial slices, 96×96 mm field of view, 200 volumes per run) because a malfunction of the 8-channel coil meant that data were acquired with a single-channel coil that required a slower TR to achieve the same resolution.

All imaging parameters and analyses were otherwise similar to Experiment 7, with the exception that training examples were not averaged across the first and second halves of the experiment since the smaller number of runs in this experiment (12 v. 44) would have produced unacceptably small degrees of freedom for inference.

Models for ROI-defining univariate analyses were fit to smoothed volumes ($\sigma = 2$, 4.7 mm FWHM) and included one regressor per block type and run plus a session mean for each run ($9 \times 44 + 44 = 440$ regressors in total). Contrasts were estimated based on the fitted responses.

5.3.2 Results

5.3.2.1 Classification effects in macaque lateral temporal lobe

I tested for head view effects through a searchlight classifier analysis. A bilateral anterior STS region discriminated head view (left v. direct v. right), while posterior STS effects were confined to the left hemisphere (Figure 5.5). The left posterior and right anterior effects overlapped completely with motion and category sensitivity (moving v. static and face v. object), while the left anterior STS overlapped only with category sensitivity.

5.3.2.2 Classification effects elsewhere in the brain

Outside STS, strong classifier effects were apparent in posterior temporal and occipital regions likely corresponding to visual areas V1, V2, V3 and V4 (Figure 5.6). These effects appeared left-lateralised for head view discrimination, but were bilateral for motion and category sensitivity. The three effects overlapped and much like in STS (Figure 5.5), any overlap involving head view discrimination generally involved both of the other comparisons.

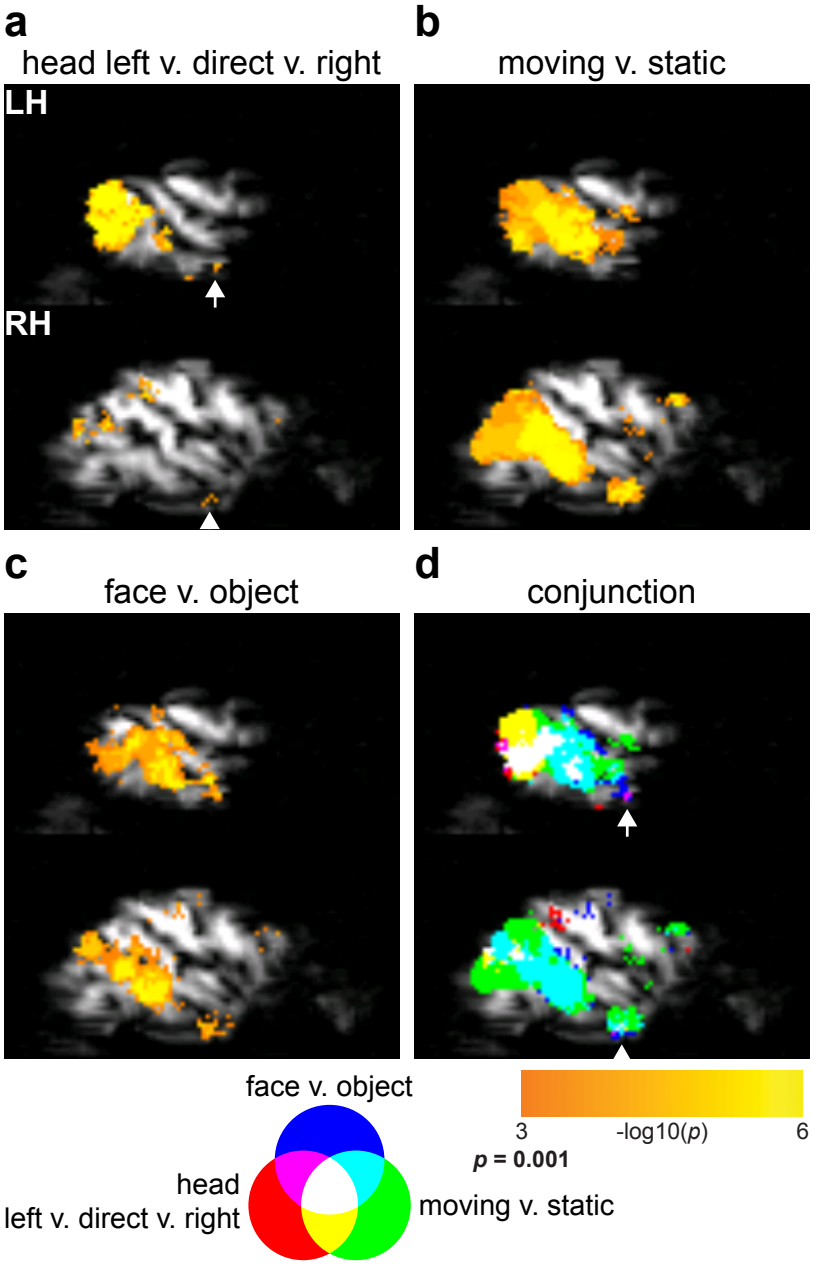


Figure 5.5: Searchlight maps of macaque lateral temporal lobe. Anterior STS effects are illustrated with white arrows. (**a**) Head view discrimination in left middle and anterior STS and in right anterior STS. (**b**) Extensive motion sensitivity along STS. (**c**) Category sensitivity in multiple STS regions. (**d**) A conjunction of the results in panels a-c shows that gaze discrimination generally overlaps with both motion and category sensitivity. The maps are thresholded at $p < 0.001$ (uncorrected).

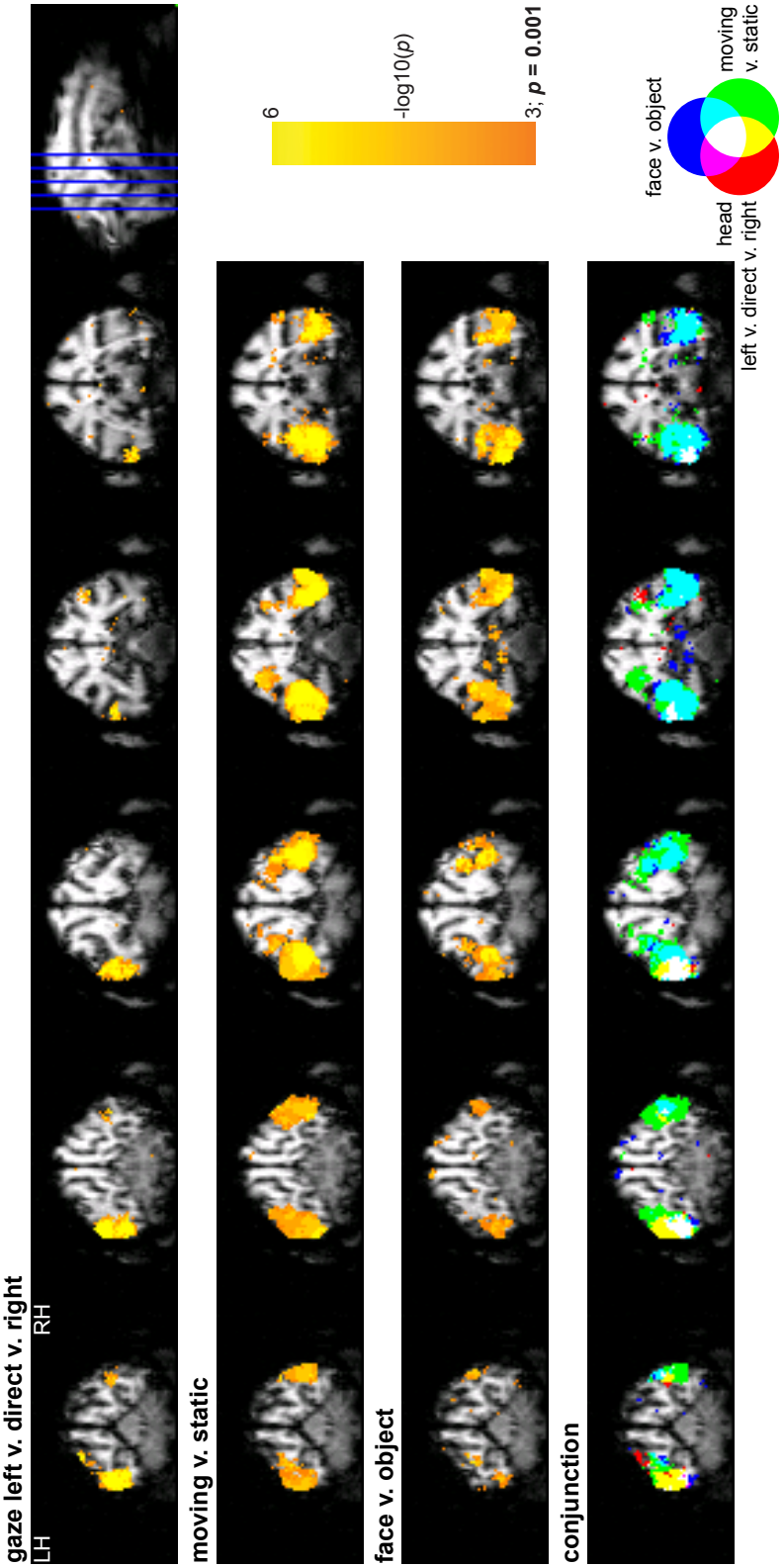


Figure 5.6: Searchlight maps of macaque visual cortex. The subject shows a strongly left-lateralised gaze discrimination effect in regions likely corresponding to early visual areas. Effects of motion and category sensitivity are more widespread and bilateral and overlap with the gaze effect. The maps are thresholded at $p < 0.001$ (uncorrected).

5.3.2.3 Region of interest analysis

I defined face-responsive patches in a smoothed univariate analysis of face blocks compared to non-face object blocks. The faces were presented in separate blocks from the head view conditions to avoid circularity. I identified 6 bilateral patches where faces elicited stronger responses than objects (Figure 5.7a). Thresholds for ROI definition varied ($p < 0.001$ to $p < 0.0001$). The patches were labelled according to previous descriptions (Tsao et al., 2003, 2008). I selected voxels inside each face patch for a subsequent functional ROI classifier analysis. This analysis revealed significant head view discrimination in left posterior lateral face patch (PL), MF and ML (Figure 5.7b). The location of left MF/ML was broadly consistent with the middle STS effect in the searchlight analysis (Figure 5.5, left panel), while the PL effect may reflect voxels from the large cluster of effects in the more posterior lunate sulcus (Figure 5.6, top panel) likely corresponding to area V4. Note that the previously discussed bilateral anterior STS head view effects in the searchlight analysis were not reflected by significant effects in any of the anterior temporal face patches, even though the anterior STS searchlight effects fell inside the AL face patch in both hemispheres (cf. Figures 5.5 and 5.7a).

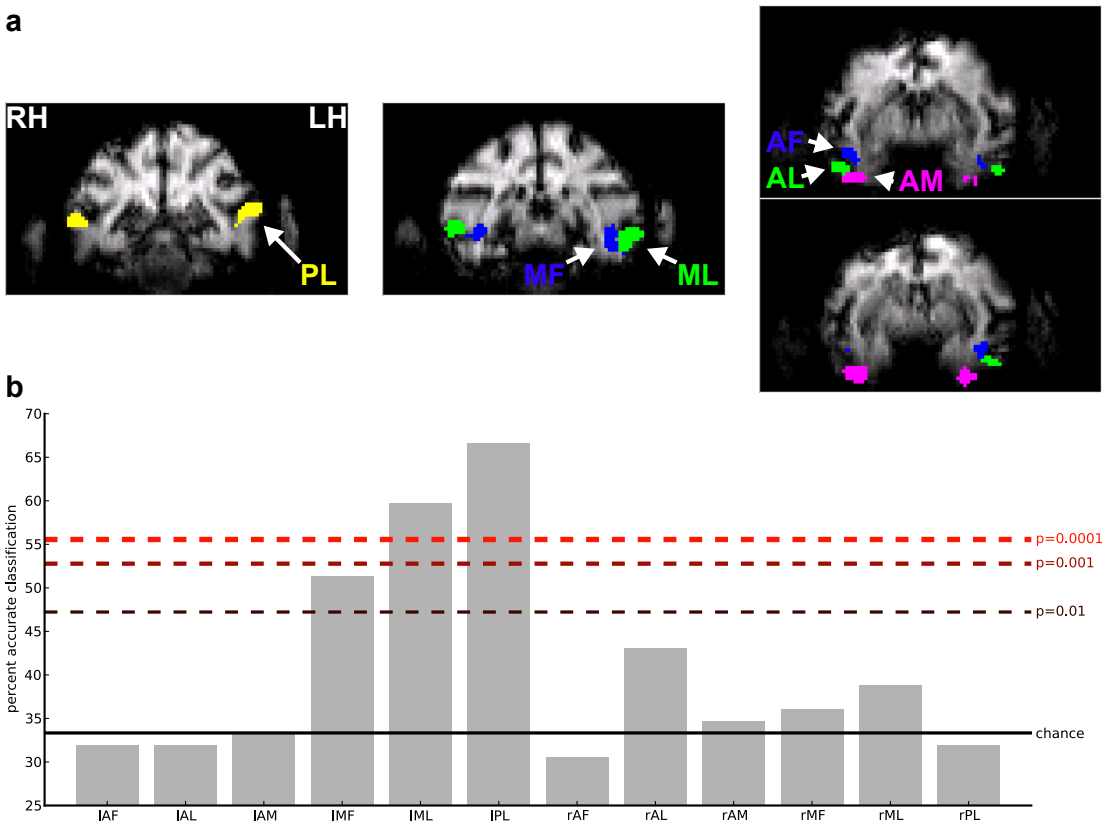


Figure 5.7: Region of interest analysis of head view discrimination inside face-responsive patches. (**a**) Face patches were defined using independent data. Six patches were identified in each hemisphere of the temporal lobe. The patches are shown with labels according to conventions in the field (Tsao et al., 2008). (**b**) Head view discrimination inside the face patches. Above-chance classification was observed in three left-hemisphere patches: MF, ML and PL.

5.4 Comparison of MVPA and univariate effects across Experiments 7-8

The three macaque fMRI datasets that constitute Experiments 7-8 provide an opportunity to assess whether MVPA methods reveal information that is not present in conventional mass-univariate contrasts applied to the same data. This comparison is particularly relevant since Experiment 1 (Chapter 3) showed that direction-sensitive responses to dynamic head turns in human anterior STS/STG only appear in searchlight MVPA and not in univariate contrasts (Section 3.2.2.4). Are similar advantages for MVPA in revealing direction-sensitive responses to social cues apparent in macaque fMRI data?

5.4.1 Methods

The multivariate analysis was based on the same MVPA methods that were described previously in this chapter. I also carried out a standard univariate analysis of the same data and computed an F contrast that encapsulates any pairwise difference between the 3 gaze or head view directions (Formula 5.2; Note that the third pair $[1, 0, -1]$ is redundant as it is a linear combination of the previous pairs). I generated p maps from the resulting F statistics to allow direct comparison between the p values obtained for univariate and multivariate methods.

$$Fcontrast = \begin{array}{c} \begin{array}{ccc} left & direct & right \end{array} \\ \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \end{array} \tag{5.2}$$

Univariate analysis is typically carried out on smoothed data while MVPA is typically carried out on unsmoothed data. In order to determine how smoothness affected any differences between univariate and multivariate results I repeated all analyses on smoothed and unsmoothed data ($\sigma = 2$, 4.7 mm FWHM).

5.5 Results

5.5.1 Differences between univariate and multivariate effects in macaque lateral temporal lobe

Figures 5.8, 5.9 and 5.10 show comparisons between univariate contrasts and searchlight MVPA for the three macaque fMRI datasets. Here I focus on effects in anterior STS, in line with the previously discussed evidence that this region may be important in discriminating the direction of social cues. Although the effects of social cue direction in anterior STS appeared consistently across % hemispheres in multivariate analysis of unsmoothed data, effects in anterior STS were less consistent when smoothing was applied or when univariate contrasts were computed in place of the searchlight MVPA. Smoothing reduced the multivariate effects to ⅘ hemispheres and the same proportion of hemispheres showed effects in anterior STS for smoothed and unsmoothed univariate contrasts. These findings suggest that although neither MVPA nor unsmoothed data is necessary for identifying direction-specific effects of social cues in anterior STS, the particular combination of MVPA and unsmoothed data may allow these these effects to be identified more consistently.

The comparison between univariate and multivariate analysis also suggests the presence of between-subject differences in how well each method performs. For example, subject s02 in the gaze direction experiment showed strong direction-specific effects specifically in the smoothed univariate analysis (Figure 5.9d). No comparable advantage for the combination of univariate analysis of smoothed data was evident in the other two datasets (Figures 5.8, 5.10). Such advantages for particular combinations of analysis method and smoothed or unsmoothed data appear to be idiosyncratic to each dataset in this small sample. Although multivariate analysis of unsmoothed data identifies direction-specific effects of social cues most consistently across datasets, these findings suggest that further advantages may be had by selecting the method based on the individual dataset. However, such selection must be carried out on independent training data to avoid biased inferences.

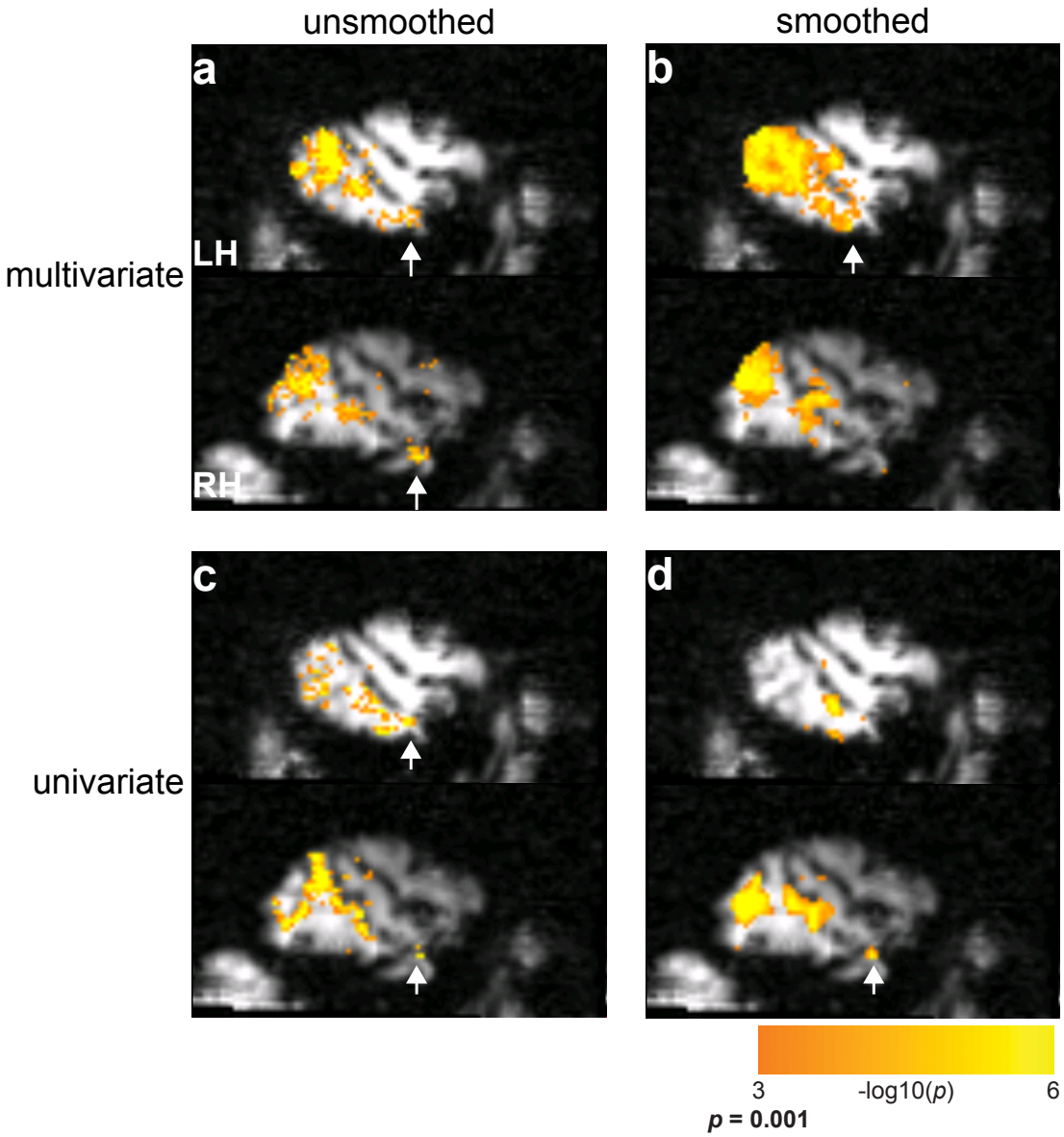


Figure 5.8: Gaze direction effects in macaque lateral temporal lobe as revealed by searchlight MVPA (**a-b**) or a comparable univariate F contrast (**c-d**) of unsmoothed (**a,c**) or smoothed (**b,d**) data (subject s01). Effects in anterior STS are illustrated with white arrows. It can be seen that both univariate and multivariate analyses produce bilateral anterior STS effects in unsmoothed data. The effects of smoothing are inconsistent in this subject. Multivariate effects in left anterior STS remain after smoothing, while effects in right anterior STS are reduced to non-significance. The opposite pattern of effects occurs for the univariate analysis: smoothing reduces left anterior STS effects to non-significance while right anterior STS effects are unaffected. The maps are thresholded at $p < 0.001$ (uncorrected).

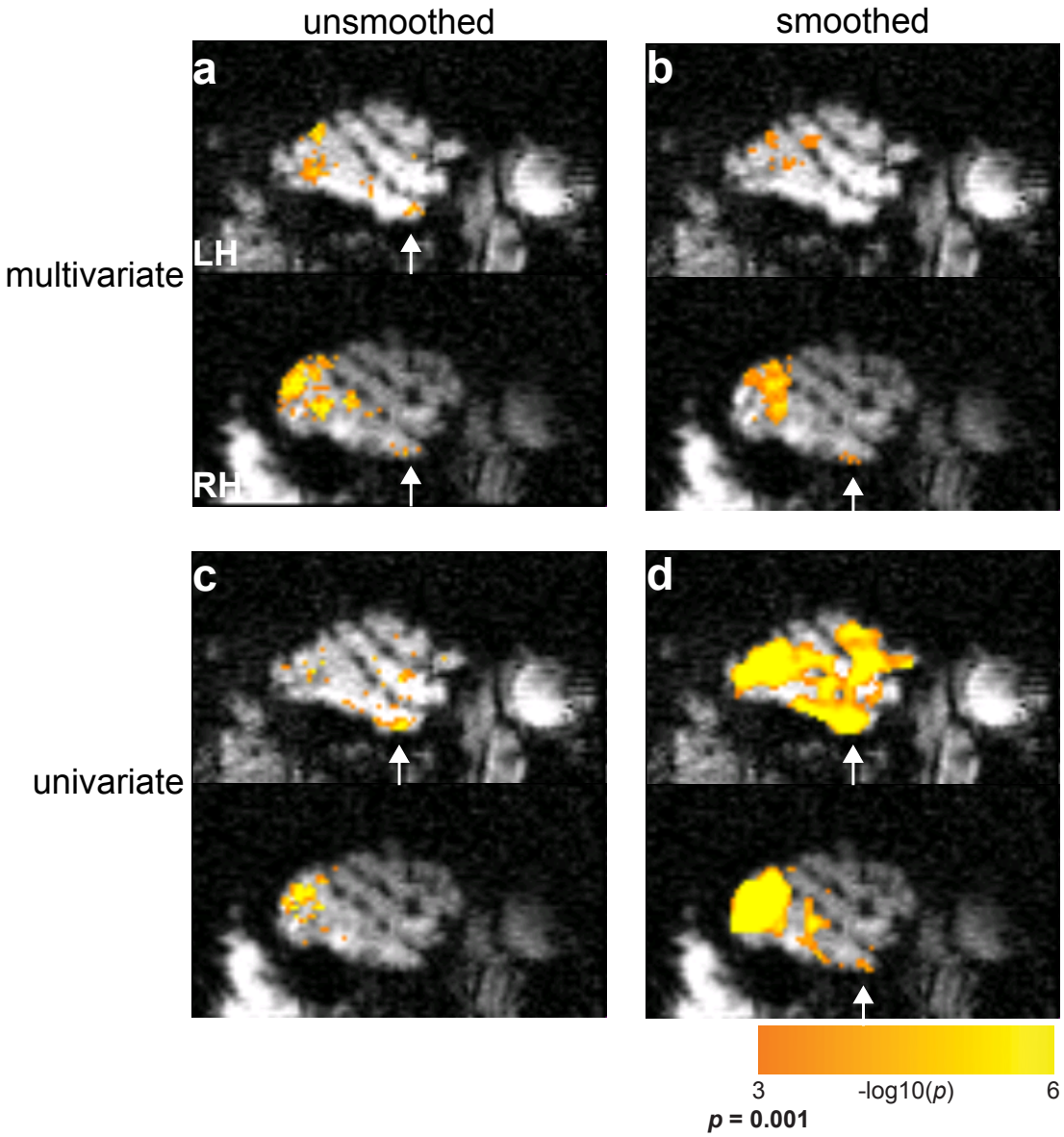


Figure 5.9: Gaze direction effects in macaque lateral temporal lobe as revealed by searchlight MVPA (**a-b**) or a comparable univariate F contrast (**c-d**) of unsmoothed (**a,c**) or smoothed (**b,d**) data (subject s02). Effects in anterior STS are illustrated with white arrows. It can be seen that the multivariate analysis produces bilateral anterior STS effects in unsmoothed data, while the univariate analysis produces left-lateralised effects. Smoothing enhances the univariate results, with strong bilateral effects in anterior STS. By contrast, smoothing reduces the multivariate effect in left anterior STS to non-significance and does not enhance the effect in right anterior STS. The maps are thresholded at $p < 0.001$ (uncorrected).

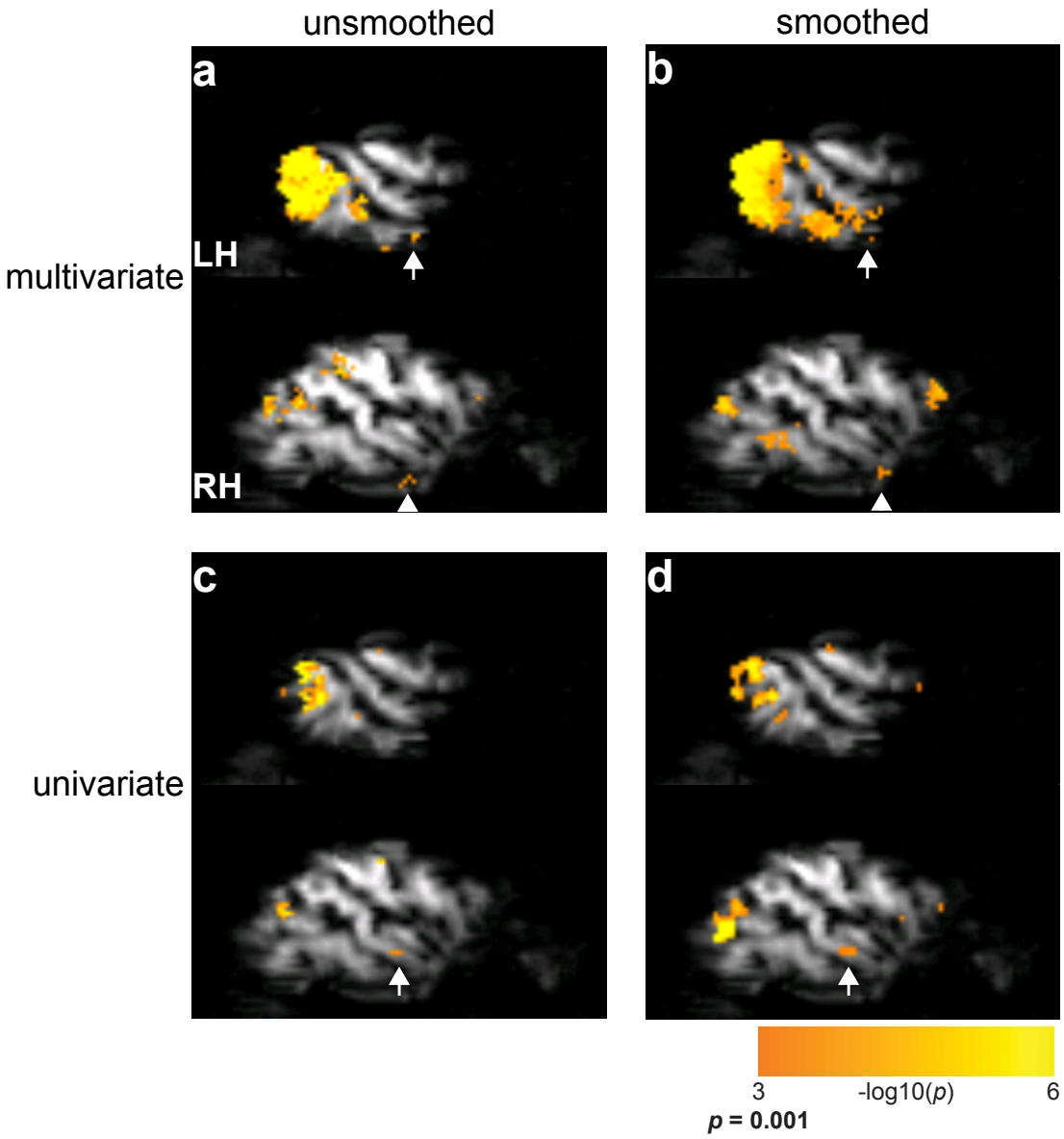


Figure 5.10: Head view effects in macaque lateral temporal lobe as revealed by searchlight MVPA (**a-b**) or a comparable univariate F contrast (**c-d**) of unsmoothed (**a,c**) or smoothed (**b,d**) data (subject s01). Effects in anterior STS are illustrated with white arrows. It can be seen that regardless of whether the data is smoothed before analysis, the multivariate analysis produces bilateral anterior STS effects, while the univariate analysis produces right-lateralised effects. The maps are thresholded at $p < 0.001$ (uncorrected).

5.5.2 Whole-brain comparison between univariate and multivariate effects

In order to investigate the relationship between univariate and multivariate effects I also tested how the strength of these effects compared in each voxel of the full volume in the 3 datasets (Figure 5.11). This comparison revealed several differences between the effects obtained through univariate and multivariate analysis. In all 3 unsmoothed datasets, a majority of voxels showed stronger effects in multivariate than in univariate analysis (Figure 5.11a,c,e). A similar benefit for multivariate analysis was evident for two of the smoothed datasets (Figure 5.11b,f). However, the remaining dataset showed a strong reversal in favour of univariate analysis after smoothing was applied (Figure 5.11d).

Even though multivariate analysis tended to produce stronger effects in most voxels, the distribution of voxel effects in all plots is wider on the univariate than the multivariate axis, suggesting that the strongest individual effects appeared in the univariate analysis. Thus, the benefit of multivariate analysis in these datasets appears to lie in detecting weaker effects that did not reach significance in the univariate analysis. This pattern was more pronounced in the analyses of smoothed data where the peaks of the univariate distribution became stronger while the multivariate effects were unchanged or reduced.

For both univariate and multivariate analysis, the strongest effects were often found far from the slope line which indicates that strong effects in one analysis did not necessarily coincide with strong or even statistically significant effects in the other analysis. Thus, although multivariate analysis generally outperformed univariate analysis, each analysis identified distinct voxels.

An important limitation to the comparison between univariate contrasts and searchlight MVPA is that the multivariate effect at a given voxel is a summary of the voxel pattern effect across all voxels inside the searchlight sphere, while the univariate effect is calculated independently for each voxel. Consequently, searchlight maps appear smoother than univariate F or T maps of the same data. Such smoothing differences may represent a more parsimonious account of apparent differences between the effects captured by univariate and multivariate analysis (cf. Figure 5.11). To test this possibility, I carried out another comparison, this time between unsmoothed multivariate

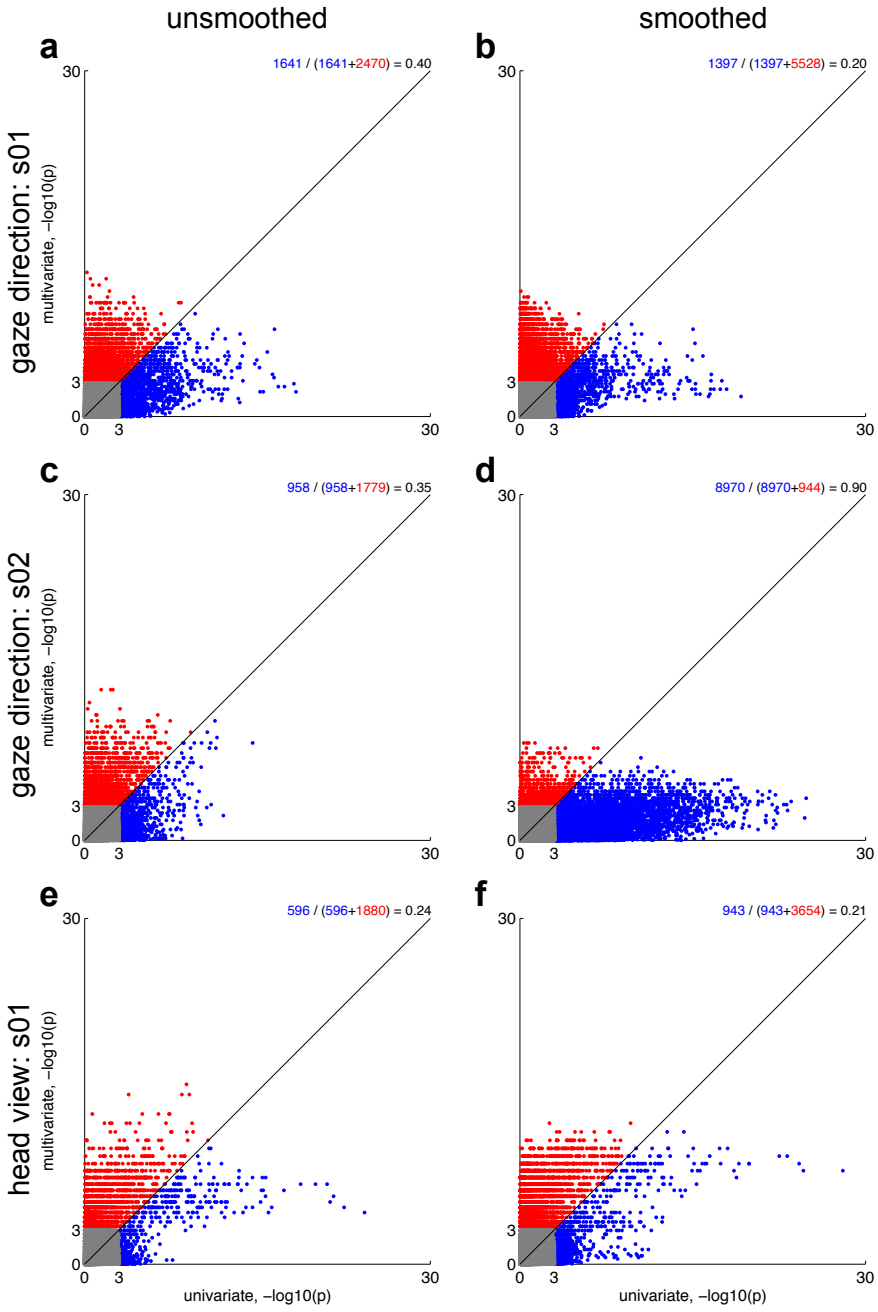


Figure 5.11: Voxel-wise plots of the effects of searchlight MVPA (vertical axis) against univariate F contrasts (horizontal axis). Effects in unsmoothed (**a,c,e**) and smoothed (**b,d,f**) data are shown for 3 datasets: gaze direction in two subjects (**a-d**) and head view in one subject (**e-f**). Voxels that exceed a significance threshold of $p < 0.001$ (uncorrected) for either analysis are colour-coded according to whether the multivariate (red) or univariate (blue) produces stronger effects (above/below the slope line). The proportion of significant voxels where the univariate analysis outperforms the multivariate analysis is summarised by the index in the top right corner of each panel where 0 reflects stronger effects in the multivariate analysis at all significant voxels and 1 reflects stronger effects in the univariate analysis at all significant voxels.

analysis and smoothed univariate analysis (Figure 5.12). This comparison between unsmoothed multivariate and smoothed univariate analyses produced qualitatively similar results as when both analyses were applied to smoothed data (Figure 5.11b,e,f). As in the previous comparison, the strongest effects for a particular analysis often appeared away from the slope line. Thus, univariate and multivariate analyses identified distinct voxels even when the relative smoothness was coarsely matched across analyses.

5.6 General discussion

The experiments in this chapter were designed to test whether MVPA of macaque fMRI data can be used to distinguish response patterns evoked by different gaze directions or head views in macaque STS. These experiments are in a position to act as the 'missing link' between previous macaque electrophysiology evidence that single cells in STS are tuned to head view and gaze direction (Perrett et al., 1992, 1985; De Souza et al., 2005), and the human fMRI studies reported previously in this thesis. Across Experiment 7 (gaze direction) and Experiment 8 (head view), all monkeys showed direction-specific coding of social cues in both hemispheres of anterior STS. Although these effects were moderate and appeared in a small sample of subjects, they are consistent both with what is known from macaque electrophysiology and with the results of the previous human fMRI experiments (Experiments 1,5). These experiments provide initial evidence for two fundamental levels of correspondence. First, human and macaque representations of head view and gaze direction in anterior STS appear to be homologous when measured with similar methods. Second, single cell data and fMRI responses in macaques lead to similar proposed roles for anterior STS (see also Tsao et al., 2006). This combination of anatomical homology across humans and macaques and methodological consistency across electrophysiology and fMRI demonstrate the validity of the central approach in this dissertation: single unit evidence from macaque STS can plausibly be used to generate hypotheses for human fMRI experiments.

5.6.1 Effects outside anterior STS

Beyond anterior STS, I also observed strong gaze and head view discrimination effects in regions of posterior visual and temporal cortex and in PL in posterior STS. With

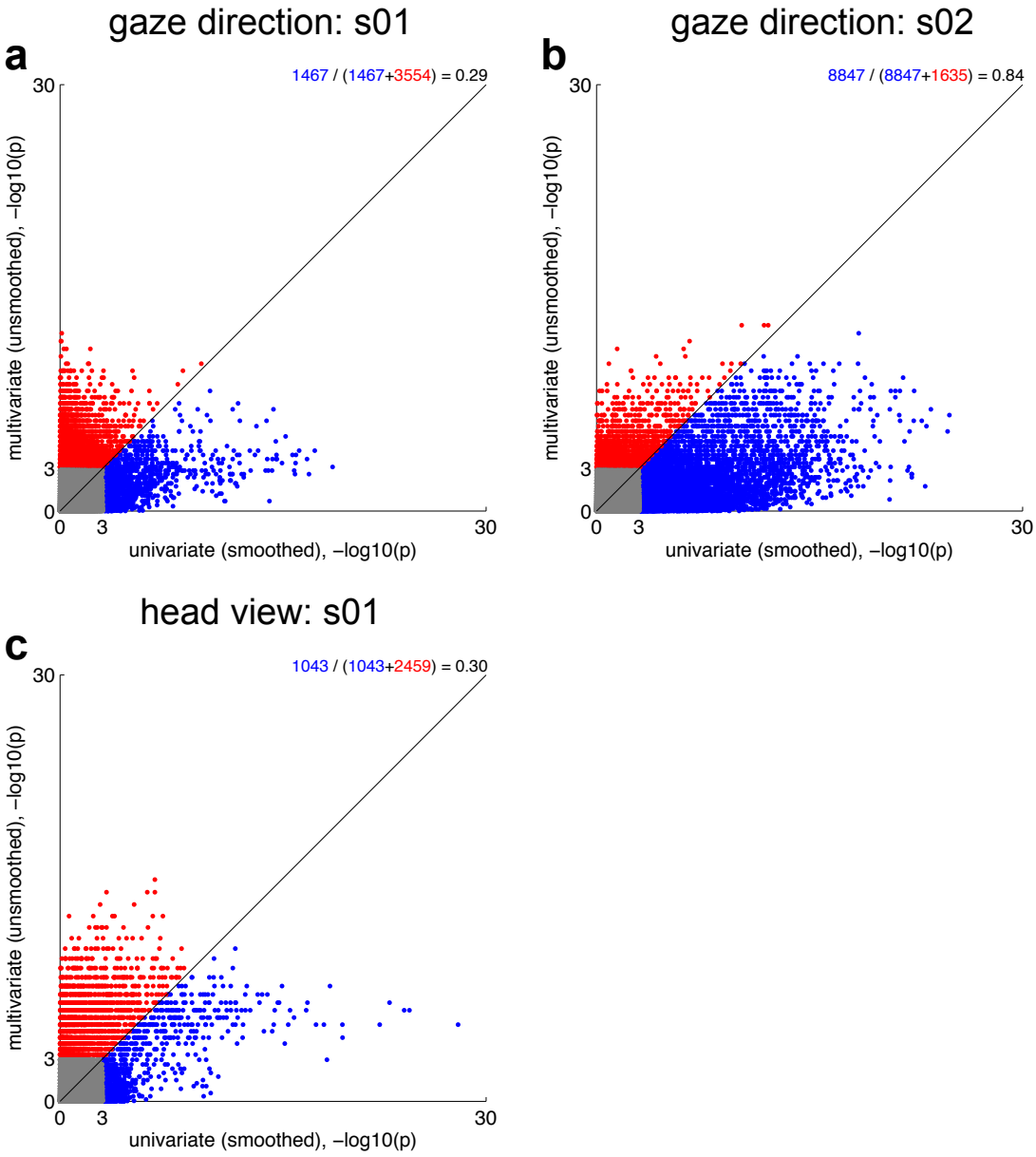


Figure 5.12: Voxel-wise plots of the effects of searchlight MVPA applied to unsmoothed data (vertical axis) against univariate F contrasts applied to smoothed data (horizontal axis). Effects are shown for 3 datasets: gaze direction in two subjects (**a-b**) and head view in one subject (**c**). Voxels that exceed a significance threshold of $p < 0.001$ (uncorrected) for either analysis are colour-coded according to whether the multivariate (red) or univariate (blue) produces stronger effects (above/below the slope line). The proportion of significant voxels where the univariate analysis outperforms the multivariate analysis is summarised by the index in the top right corner of each panel, where 0 reflects stronger effects in the multivariate analysis at all voxels and 1 reflects stronger effects in the univariate analysis at all voxels.

the exception of posterior STS, there are no reports of gaze direction or head view cells in these regions, but these findings are consistent with the head turn classification effect that I observed in human primary visual cortex in Experiment 1 and 4 (Chapter 3). It is important to note that as in in Experiment 4, these effects cannot be explained by eye movements since the subjects were fixating. However, in light of previous evidence that humans and macaques are both spatially cued by perceived gaze (Friesen & Kingstone, 1998; Deaner & Platt, 2003), one possibility is that these effects reflect attentional re-orienting.

Experiments 7-8 share an important limitation in that the different social cue directions are physically different and would thus produce slightly different response patterns in any retinotopically-organised region. The overlap with motion sensitivity, a known property of early visual regions (Mikami, Newsome & Wurtz, 1986b; Huk et al., 2002; Seymour et al., 2009), is consistent with this retinotopic account. Further studies are needed to disentangle macaque fMRI responses to the direction of another's attention from responses to the physical changes that convey this direction (cf. Experiment 5, Chapter 4). Such studies are particularly urgent in light of the finding that motion sensitivity coincides with gaze direction and head view sensitivity in macaque posterior STS. It is currently unclear if this overlap reflects distributed representations for social cues and motion processing or distinct sub-regions at spatial scales beyond what the searchlight method can resolve.

5.6.2 Relationship to face selectivity

Experiment 8 provided some initial data on overlap between head view discrimination and preferential univariate responses to faces relative to other categories. As in the human results in Experiment 1 (Chapter 3), discrimination of a socially-relevant within-face stimulus feature in anterior STS could not be related to functionally-defined face-responsive regions from a univariate analysis. This failure of the functional localiser approach also in macaques suggests that representations of head view and gaze direction may not align perfectly with the previously-reported macaque face patch system (Tsao et al., 2003, 2008). However, the current chapter does not allow strong conclusions concerning this issue as Experiment 8 was the only experiment to include a meaningful functional localiser and this experiment was based on a single subject with

relatively limited amounts of data.

5.6.3 Comparisons between univariate and multivariate analysis

Although MVPA methods are becoming a standard tool in human fMRI, this method has only rarely been applied to macaque fMRI data (for one example, see Tsao et al., 2003). It cannot be assumed that the reported advantages of MVPA in human fMRI will necessarily transfer to macaque studies, because macaque and human fMRI protocols differ on many parameters including the use of the MION contrast agent, higher resolution acqusition, smaller field of view and anatomical differences in brain size. The comparisons reported here show that similar to the human data (Experiment 1, Chapter 3), searchlight MVPA identifies direction-specific responses to social cues in anterior STS more consistently than univariate contrasts. However, univariate contrasts also identified direction-specific anterior STS responses, although less consistently. Indeed, in one subject univariate contrasts on smoothed data outperformed MVPA. These findings suggest that although MVPA methods appear to be applicable to macaque fMRI, the advantages over univariate contrasts may not be as striking as for human fMRI. The comparisons here cannot address the origins of this advantage. In particular, it is possible that MVPA methods benefit less from the signal-to-noise benefits of the MION contrast agent relative to univariate analysis but this account requires tests with BOLD macaque fMRI.

5.6.4 Conclusions

This chapter describes experiments that provide the first evidence that the direction of social cues can be distinguished based on anterior STS response patterns in macaque fMRI data. These findings provide a crucial link between previous macaque neurophysiology data from this region and the human fMRI experiments described in previous chapters. The proposed role for anterior STS in distinguishing the direction of social cues is thus supported by converging evidence from multiple species.

Chapter 6

Discussion

This thesis describes experiments in which I used MVPA of fMRI data from humans and macaques to understand how the direction of visual social cues, including head turns, gaze direction and head view, are represented in the brain. In this chapter I summarise the main findings that emerged from this research and identify common patterns across the studies. I also discuss limitations of the thesis before outlining directions for future research.

6.1 Summary of results

Humans are exceedingly sensitive to social cues, including gaze direction and head turn. Although single cells in macaque anterior STS are tuned to the direction of such social cues, neuroimaging studies have produced somewhat inconsistent evidence concerning how human STS codes social cues. Most human neuroimaging studies that were based on conventional univariate contrasts found involvement of posterior rather than anterior STS (Nummenmaa & Calder, 2009), while one fMRI adaptation study implicated the presence of gaze direction sensitivity in anterior STS instead (Calder et al., 2007). By applying MVPA to this question, I obtained evidence that consistently implicated anterior STS in coding the direction of social cues in both human and macaque fMRI. This result is consistent with both previous macaque electrophysiology and human fMRI adaptation studies.

In Experiment 1, classifier MVPA methods revealed that a region in right anterior

STS/STG distinguishes the direction of dynamic head turns. I also found that direction discrimination in this region does not reflect a more general role in coding rotational movement: the direction of ellipsoid rotation could not be discriminated in this region and discrimination of head turn direction was significantly better than discrimination of ellipsoid rotation. Although a follow-up eye tracking study suggested the presence of eye movement confounds (Experiment 2), I was able to revise the design of the experiment to control eye movements (Experiment 3). A follow-up fMRI experiment showed that head turn discrimination in anterior STS remained even in the absence of significant eye movement effects (Experiment 4). These findings suggest that anterior STS/STG discriminates the direction of dynamic head turns, consistent with previous evidence that anterior STS shows direction-specific adaptation to gaze direction (Calder et al., 2007).

In Experiment 5, I sought to investigate the representational content of gaze direction representations in STS. I found that a region in anterior STS responded to gaze direction in a manner that was invariant to head view and physical image features. This finding suggests that part of the anterior STS representation of social cues is coded in terms of a high-level representation of social attention direction, rather than coding the physical features of eye position or head view as such. However, previous behavioural gaze discrimination studies suggest that head-gaze integration is imperfect, with the perceived gaze direction being shifted in the opposite direction of the head view (Gibson & Pick, 1963). To address this issue, Experiment 6 was designed to collect behavioural discrimination data for the same stimuli as in Experiment 5. The results showed that although head view elicited a small effect on gaze discrimination, performance was well accounted for by a view-invariant gaze direction predictor that assumed perfect head-gaze integration. Furthermore, the results in Experiment 5 were highly similar when each individual volunteer's fMRI data was fitted against their own gaze discrimination performance from Experiment 6. Thus, both behavioural and fMRI representations of gaze direction could be explained well without accounting for imperfect head-gaze integration or volunteer-specific gaze discrimination effects.

In Experiments 7-8, I applied similar fMRI analysis methods that I previously used with humans fMRI to study gaze direction and head view representations in macaque STS. These experiments addressed whether human and macaque representations of social cues are comparable when the same neuroscientific measure is used. Although

these experiments were preliminary, the results suggest that both gaze direction (Experiment 7) and head view (Experiment 8) can be distinguished in macaque anterior STS. These results match both the previous human fMRI experiments in this thesis (Experiments 1,5) and previous literature from macaque electrophysiology (Perrett et al., 1985, 1992; De Souza et al., 2005).

6.2 Advances of the current thesis

This thesis provides new insights into how the direction of social cues is represented in the human and macaque brain. These advances were achieved by analysing fMRI data with classification- and RSA-based forms of MVPA. Here I discuss some of the key empirical findings for particular brain regions and the developments in MVPA methods that support the findings.

6.2.1 Neural representations of social cues in human and macaque

6.2.1.1 Role of anterior STS

Direction-specific responses to social cues were evident in anterior STS in all described experiments, whether the social cues were dynamic head turns (Experiment 1) or static gaze cues across different head views (Experiment 5) in human data, or gaze and head view cues in the macaque data (Experiments 7-8). This contrasts with previous human neuroimaging evidence, which had mostly concerned distinctions between direct and averted gaze in posterior rather than anterior STS (Nummenmaa & Calder, 2009). The consistent anterior STS effects reported here suggest that when MVPA methods are used, there is no conflict between the results of human/macaque fMRI studies and previous macaque electrophysiology research: anterior STS is implicated in all cases. These findings are also consistent with previous fMRI adaptation evidence for direction-specific gaze representations in anterior STS (Calder et al., 2007), thus providing convergent evidence across distinct fMRI methods.

It is less clear how regions that discriminated gaze direction (Experiments 5,7) correspond to regions that discriminated dynamic head turns or static head views (Experiments 1,8). In humans, the head turn discrimination effects in Experiment 1 were posterior and superior to the gaze discrimination effects in Experiment 5. In macaques,

effects of gaze direction (Experiment 7) and head view (Experiment 8) appeared to engage similar anterior STS regions, but conclusive comparisons cannot be made as only one macaque subject participated in Experiment 8. There is evidence that neurons responsive to head turns in macaque anterior STS are unresponsive to static head view cues (Hasselmo et al., 1989; Perrett et al., 1985), but it remains unclear whether such dynamic head turn neurons are spatially segregated from neurons that also respond to static views.

6.2.1.2 Role of posterior STS

Effects in posterior STS were less consistent than in anterior STS. Human posterior STS did not discriminate head turn direction in Experiment 1 even at reduced thresholds. In Experiment 5, posterior STS responded to static gaze direction, but these effects were reduced when the influence of head view or physical image features was removed. The results in Experiment 5 appeared broadly consistent with a processing hierarchy for social cues in STS, where representations of gaze direction in posterior STS remain influenced by head view, while anterior regions respond to gaze in a view-invariant manner. Although this proposal fits with electrophysiological results demonstrating how view-invariant representations of face identity emerge in macaque temporal lobe (Freiwald & Tsao, 2010), further investigation is needed to test this explanation for social cue representations.

Although macaque posterior STS also showed direction-specific responses to social cues, including gaze direction (Experiment 7) and head view (Experiment 8), interpretation of these results is complicated by the fact that macaque posterior STS also contains motion area MT and the adjacent lunate sulcus contains area V4. Indeed, direction-specific gaze direction and head view effects in macaque posterior STS often appeared to originate in stronger effects in lunate sulcus. Experiments 7-8 did not include retinotopic localisers so it is unclear at present whether direction-specific social cue responses in the vicinity of posterior STS can be related to retinotopic biases detected by these areas or, alternatively, to potentially face-specific coding in PL. Previous evidence that MT appears in different posterior temporal regions in humans and macaques (Orban et al., 2004) also makes it difficult to conduct direct comparisons between macaque and human social cue representations in this region. Thus, the cur-

rent results are not conclusive regarding the homology of social cue representations in macaque and human posterior STS, although there is initial evidence that anterior STS representations of social cues are similar across these species.

6.2.1.3 Role of brain regions outside STS

Across all experiments, the most consistent effects outside STS appeared in retinotopic visual cortex. Regions likely corresponding to V1-V4 discriminated head turn direction in Experiments 1 and 4 (human volunteers). Similarly, regions in macaque visual cortex discriminated gaze direction (Experiment 7) and head view (Experiment 8).

Do these responses reflect representations of the direction of social cues in early visual areas? There is no macaque electrophysiological data describing how single cells in these regions respond to social cues, but several interpretations of the fMRI findings are nevertheless possible. First, responses may reflect attentional cueing since social cues elicit shifts in spatial attention (Friesen & Kingstone, 1998) and fMRI responses in early visual cortex are modulated by attention (Kastner, Pinsk, Weerd, Desimone & Ungerleider, 1999). A second related possibility is that fMRI responses in early visual cortex reflect feedback from social cue representations in STS. Such feedback may serve to enhance visual processing of the cue itself to enable fine-grained direction discrimination, or to mediate shifts in spatial attention in response to the perceived cue direction. Finally, the direction-specific social cue effects in early visual cortex may occur because the retinotopic organisation of early visual cortex renders these regions sensitive to physical differences between the different social cue directions in the experiments (see also Section 6.3.1.2). The last of these interpretations is made likely by the fact that no early visual effects appeared in Experiment 5, which tested for view-invariant representations of gaze direction by also manipulating head view. This design de-couples physical differences from social cue direction, so the fact that this was the only experiment where early visual effects were not found may indicate that such effects in the other experiments were driven by retinotopic sensitivity to low-level physical differences between the social cue directions.

I also observed strong view-invariant gaze direction effects in precuneus in Experiment 5. No comparable effects were observed in humans in response to head turns (Experiment 1), or in the macaque gaze direction or head view studies (Experiments

7-8). It is possible that these effects reflect attentional re-orienting (Cavanna & Trimble, 2006; Laube et al., 2011), although the role of precuneus in gaze coding remains a topic for further study.

6.2.2 Methodological developments

6.2.2.1 Benefits of MVPA relative to other methods

I made comparisons between standard univariate contrasts and searchlight MVPA in Experiments 1, 7 and 8. Although MVPA effects sometimes coincided with univariate effects at similar statistical thresholds, effects in anterior STS were only consistently observed with MVPA. These findings demonstrate the utility of MVPA in revealing STS effects of social cue direction, although it must be acknowledged that fMRI adaptation methods have also been used to reveal direction-specific representations of gaze direction (Calder et al., 2007). The results from the current thesis complement these fMRI adaptation findings by demonstrating that direction-specific anterior STS responses to social cues can be detected also with more flexible MVPA methods. The neural mechanisms that support MVPA and fMRI adaptation are both poorly understood (for critical examples, see Freeman et al., 2011; Krekelberg, Boynton & Vanwezel, 2006). The fact that both methods produce congruent results is reassuring since this makes it unlikely that direction-specific social cue responses in anterior STS occur as an artefact caused by the idiosyncrasies of the chosen analysis method.

6.2.2.2 Developments in representational similarity analysis

Experiment 5 (Chapter 4, see Appendix B for paper) is the first published account of how RSA can be combined with searchlight mapping. Previous accounts of RSA primarily emphasised its use in exploratory analysis of functional and anatomical regions of interest (Kriegeskorte et al., 2008). Experiment 5 demonstrates that RSA is also a powerful tool for hypothesis testing and whole-brain mapping. Unlike classifier-based MVPA where the data must be split during crossvalidation, RSA can be carried out on the complete dataset which maximises detection power.

It is relatively straight-forward to localise regions where a particular predictor RDM explains the data RDM. In the searchlight framework, the Spearman rank corre-

lation that summarises the strength of the predictor-data relationship is mapped back to the centre voxel, much like classification accuracy maps are generated by mapping percent correct classification back to the centre voxel of each searchlight. These r maps can then be normalised, smoothed and entered into permutation-based group analyses. Experiment 5 also showed how Spearman partial correlation can be used to generate searchlight maps that summarise how well a given predictor-data relationship holds when the influence of a competing predictor is removed. Previous studies used Spearman r as a metric for how well each predictor RDM fits the data and compared r across predictor-data pairs (see e.g., Figure 8 in Kriegeskorte et al., 2008). This approach is useful for exploratory purposes, but it does not consider dependencies between correlated predictors. The aim of the predictor comparisons in Experiment 5 was to show that the experimental effect of view-invariant gaze direction could not be explained by competing predictors that coded the stimuli in terms of head view and physical image features, even though these competing predictors correlated with view-invariant gaze direction. Spearman partial correlation was chosen for this analysis because it accounts for correlations between predictor RDMs and, as a rank measure, it assumes only a monotonic relationship between predictor and data RDMs. This analysis showed that the effect of view-invariant gaze direction in anterior STS survived removing the influence of head view or physical image features. Given the difficulties in fully removing physical confounds from comparisons between different social cue directions (I return to this point in Section 6.3.1.2), this partial correlation RSA approach may prove useful in future studies.

6.3 Limitations of the current thesis

Beyond the empirical findings reported in the experimental chapters, this thesis has also been a learning experience for its author. In this section I describe some of the design and method limitations encountered during this work and how such issues may be avoided in future studies.

6.3.1 Experimental design limitations

6.3.1.1 Comparing social and non-social movement

Experiments 1 and 2 used a comparison between videos of real-life turning heads and computer generated rotating ellipsoids. A region in anterior STS/STG was found to discriminate head turn direction, but not direction of motion in rotating ellipsoids. These findings indicate that discrimination of head turn direction in anterior STS/STG is not reducible to discrimination of any rotational motion, whether the rotating object is a head or ellipsoid. However, the strength of this demonstration of selectivity depends critically on how closely matched the head turns and ellipsoid rotations were in low-level physical features. The ellipsoids were texture mapped with the scrambled face texture from the head turns, and the motion trajectory of the rotation was matched to the kinematics of the head turn. Nevertheless, clear physical differences remain between the two stimulus types.

Physical matching for dynamic stimuli is a difficult problem because if the physical features of each individual frame is matched across video pairs, for instance by scrambling each video frame separately, this disrupts the optic flow that yields the motion percept across frames. Alternatively, it is possible to generate a single scrambled texture that then moves across frames in a similar profile to the experimental condition (similar to the approach in Experiments 1-2). However, in this case the pixel-wise matching between the two experiment and control conditions will likely be imperfect. Although some form of control motion is clearly necessary for any study of socially-relevant motion, physical matching in dynamic videos remains a significant challenge to this field.

6.3.1.2 Discriminating the direction of single social cues

Many of the experiments in this thesis are based on comparing fMRI responses to different directions in a single social cue such as head turns (Experiments 1-4), different gaze directions (Experiment 7) or different head views (Experiment 8). The interpretation of such comparisons is limited in that the direction of the social cue correlates with a concomitant physical change. Thus, brain responses could reflect coding of the social cue or coding of the physical change, whether this physical change corresponds to a feature as subtle as iris displacement in a front-facing head or the large-scale

retinotopic differences between the profile of a left- and right-turned head. Much of visual cortex is retinotopically organised (Brewer, Liu, Wade & Wandell, 2005; Larsson & Heeger, 2006; Arcaro, Mcmains, Singer & Kastner, 2009) and MVPA methods appear particularly sensitive to detecting information about retinal location (see e.g., Kravitz, Kriegeskorte & Baker, 2010; Schwarzlose, Swisher, Dang & Kanwisher, 2008; Vicente-Grabovetsky, Mitchell, Carlin & Cusack, 2011). Thus, it cannot be ruled out that the basis of direction discrimination in these particular experiments is the retinotopic differences between the two directions rather than a representation of social cues as such. These concerns are perhaps less critical for the interpretation of anterior STS effects, since there are no previous reports that this region contains retinotopic visual field representations in humans or macaques. However, retinotopic confounds become a serious concern when direction-specific effects are observed in regions known to contain retinotopic organisation, including regions corresponding to human or macaque areas V1-4 and in particular area MT, which is located in posterior STS in macaques. Direction-sensitive responses in areas with likely retinotopic organisation was apparent in Experiment 1 and in Experiments 7-8.

Physical differences between different social cue directions can be minimised but they can never be abolished so long as only one social cue is manipulated at a time. For example, consider that a leftward direction can only be conveyed by shifting some feature leftward, whether it is head view, eye position or body posture. A better solution to this problem is to de-couple physical change and social cue direction by manipulating multiple social cues simultaneously. In Experiment 5, the direction of gaze was conveyed by multiple combinations of head view and eye position relative to the head. Because head view affects a far greater portion of the image this cue comes to dominate predictions based on physical image features, while the emergent property of the head view / eye position coupling – gaze direction – yields a different set of predictions. Although it would be challenging to fully orthogonalise the social cue direction and physical image features, the use of multiple social cue combinations allows the correlation between the social cue direction and physical image features to be sufficiently reduced so that their relative influences can be dissociated.

6.3.2 Methodological limitations

6.3.2.1 First-level inference on searchlight classification maps with binomial tests

All second-level group analyses in this thesis were computed using non-parametric permutation testing. Such tests are well-suited to the experiments I report because it is unclear whether the first-level summary statistics that are used (classification accuracy, Spearman correlation and partial correlation) conform with the distributional assumptions that support conventional parametric group analysis of fMRI data. Furthermore, permutation tests in SnPM can achieve enhanced sensitivity relative to standard parametric group analysis in SPM through the use of variance smoothing (Nichols & Hayasaka, 2003). It would have been desirable to rely strictly on permutation testing also for the first-level inferences reported in Experiments 4, 7 and 8. However, at the time of conducting this research, such searchlight permutation tests were infeasible with available computing resources. The computational demands of this test scales with the number of permutations because the entire searchlight map must be re-computed for each permutation. Given that a representative classification accuracy searchlight map in Experiment 7 took a few hours to compute with available computing resources it would be impractical to generate 10000 permutations in a reasonable space of time.

However, this limitation is likely to be removed in the near future with the rise of cloud computing services including Amazon EC2 (http://aws.amazon.com/ec2, retrieved 23 October 2011) and scientific computing implementations of this service such as Opani (http://www.opani.com, retrieved 23 October 2011). Permutation testing is an easily parallelised computational problem so with 10000 available computing engines, a full permutation test need not take more time than what computing a single searchlight map takes at present. Although previous methodological evaluations have found that binomial tests produce comparable results to permutation testing for classification accuracy maps (Pereira & Botvinick, 2011) this assumption may not hold for all datasets. Future studies should not have to assume that binomial tests are appropriate if recent advances in cloud computing continue.

6.3.2.2 Volumetric and surface-based searchlight analyses

All fMRI experiments in this thesis used searchlight mapping to identify regions showing MVPA effects and, with the exception of the macaque studies in Experiments 7-8, all experiments restricted the searchlight mapping to grey matter regions as defined by the automatic segmentation routines in SPM. Grey matter masking was motivated by previous findings that masking enhances classifier MVPA performance (Oosterhof et al., 2010), presumably because grey matter voxel responses contain more information about the conditions of the experiment than white matter voxels. Grey matter masks also contain considerably fewer voxels than conventional whole-brain analysis masks, which reduces type 2 errors from multiple comparisons correction. However, grey matter masking complicates group analysis of the individual volunteer searchlight maps since volumetric normalisation to the MNI or Talairach template may result in imperfect alignment of grey matter across volunteers. In this thesis I masked group results by a group grey matter mask created as the union of the individual volunteer grey matter masks. Although this method worked well in the current experiments, it is clear that the group grey matter mask has poor anatomical precision with regard to fine gyral and sulcal grey matter structure.

A preferable approach in future research would be to use analysis methods that operate on the grey matter surface throughout data analysis. There is evidence that searchlight mapping is more sensitive when surface-defined patches are used for mapping rather than volumetric spheres (Oosterhof et al., 2010). The outcome of this analysis is a searchlight map that is coded in surface vertices rather than in volumetric voxels. A key benefit of this approach is that group analysis can now proceed using surface-based normalisation methods as implemented in software packages such as Freesurfer (Fischl, Sereno & Tootell, 1999). This approach obviates the inherent problems in analysing grey matter voxels across volunteers with volumetric normalisation.

6.3.2.3 Anatomical definitions of STS

This thesis is primarily concerned with representations of the direction of social cues in the STS region. However, there is conflicting evidence concerning the exact site of human STS that plays a role in this function. Therefore I developed a masking approach for defining the anatomical STS region in each experiment based on the

average T_1 structural volume for the sample (Section 2.2.3.2). Although this method is sensitive to variability in STS structure across volunteer samples, it is a manual and therefore subjective approach. Volumetric toolboxes such as automated anatomical labelling (Tzourio-Mazoyer et al., 2002) do not include sulci and would in any case fail to account for STS variability across different volunteer samples. As in the preceding section, this limitation may be resolved by using surface-based methods. Freesurfer includes automatic parcellation of sulci (Fischl, van der Kouwe, Destrieux, Halgren, Segonne, Salat, Busa, Seidman, Goldstein, Kennedy, Caviness, Makris, Rosen & Dale, 2004) and because surface normalisation acts to explicitly align the sulcal structure across volunteers (Fischl et al., 1999), STS mask definition for individuals and samples can be made both automatic (and therefore less subjective) and more anatomically precise.

6.4 Future research

This thesis demonstrates how MVPA can be used to reveal direction-specific responses to social cues in fMRI data. The positive results obtained here suggest several future lines of inquiry to gain a more comprehensive and mechanistic understanding of how the brain encodes the direction of another's attention.

6.4.1 Eye contact effects

Gaze perception may reflect two distinct processes: discriminating the direction of another's attention when it is averted (triadic gaze) and discriminating between averted gaze and eye contact (dyadic gaze). These mechanisms may operate in concert during natural social behaviour (for instance, eye contact may attract the observer's attention, which enables accurate perception of subsequent averted gaze), but there is nevertheless behavioural and macaque electrophysiology evidence to suggest a distinct status for eye contact. This distinction is not reflected in most neuroimaging efforts even though there are testable predictions. A comparison between different averted gaze directions should modulate codes for triadic but not dyadic gaze. Conversely, a dyadic code should be invariant to whether eye contact was established by, for instance, a shift from the left or from the right. Future imaging studies may use these predictions to

test whether distinct brain regions in for instance STS are concerned with eye contact detection and gaze direction discrimination.

6.4.2 Contextual influences on gaze representations

Knowing the object of another's attention often involves making fine-grained distinctions between small displacements in the looker's iris. Given the difficulty of this task it is perhaps unsurprising that humans readily use any environmental short cut to disambiguate gaze direction, whether the short cut is a salient object in the environment (Lobmaier et al., 2006), or hearing one's name (Stoyanova et al., 2010). Perception of gaze direction is clearly influenced by many contextual factors. Although a number of human imaging studies have explored contextual effects on gaze responses, the studies to date have either reported that a brain region distinguishes gaze direction in the absence of contextual manipulations (e.g., Calder et al., 2007) or that a brain region distinguishes the contextual meaning of gaze across gaze directions, such as the finding that posterior STS responds more when an actor gazes away from a salient object than when they gaze towards it (Pelphrey et al., 2003). Locating MVPA effects of both context and gaze direction in the same study would provide a valuable opportunity to understand how these representations interact. For example, such studies may involve parametrically manipulating both the angle of perceived gaze and the angular position of a target. Separate regions may be implicated in coding gaze direction (e.g., anterior STS), the location of the target (e.g., early visual cortex) and the match between gaze and target (e.g., posterior STS).

6.4.3 Relationship to other aspects of face processing

This thesis is focused on perception of social cues from the face. However, the face also conveys information about other features, including personal identity, gender, ethnicity, age, attractiveness and emotional state. According to a dominant model in the field, these functions segregate into a ventral stream for coding static face features and a dorsal stream for changeable features (Haxby, Hoffman & Gobbini, 2000). Static features are viewed as particularly important for identity processing, while emotion and gaze processing are held to depend particularly on changeable features. In this way,

the proposed model is similar to a previous cognitive model that also invoked separate processing streams for face identity and emotion information (Bruce & Young, 1986).

The current results are broadly consistent with this model in that direction-specific responses to social cues were found in lateral rather than ventral temporal regions. However, it remains unclear how well separated these streams are as processing continues into the anterior temporal lobe. For the case of regions that respond preferentially to faces than to other categories, posterior STS is separated by several centimetres from other proposed face regions in FFA and inferior occipital gyrus. However, proposed face areas in anterior STS (Fox et al., 2009; Pitcher, Dilks, Saxe, Triantafyllou & Kanwisher, 2011) and in anterior IT (Pinsk et al., 2009; Rajimehr, Young & Tootell, 2009) are not as readily separated since STS is on a posterior-superior to inferior-anterior axis that places anterior STS close to anterior IT. Anterior IT was previously found to distinguish face identity (Kriegeskorte et al., 2007) so the simplest account is that the two stream structure continues into temporal lobe, with face identity representations on the ventral surface and directional social cue representations on the lateral surface. However, note that the peaks in Kriegeskorte et al. (2007) are only 20 mm separated from the anterior STS effect for view-invariant gaze direction in Experiment 5. Alternatively, theorists outside face perception have argued that anterior temporal lobe is crucial to retrieval of information about unique entities (Damasio, Tranel, Grabowski, Adolphs & Damasio, 2004; Tranel, 2006). In this view, the same region may represent both different gaze directions and different facial identities. Future studies that manipulate both face identity and social cue direction will be necessary to test whether distinct anterior temporal regions code these features.

6.4.4 Comparing electrophysiological gaze responses between humans and macaques

The human fMRI experiments in this thesis were motivated by previous evidence from macaque electrophysiology, where single cells in anterior STS are tuned to the direction of social cues (e.g., Perrett et al., 1985, 1992; De Souza et al., 2005). This comparison runs both across species and across methods. Experiments 7-8 showed that consistent with this electrophysiological evidence, macaque anterior STS is sensitive to gaze direction (Experiment 7) and head view (Experiment 8) when measured with

fMRI, thus providing some support for the validity of the use of macaque electrophysiology data to generate hypotheses for human fMRI experiments. A complimentary question is whether social cue representations in STS are also similar across species when electrophysiological measures are used. This question is difficult to address since intracranial electrophysiology recordings are rarely possible in humans. One exception to this is patients with intractable epilepsy who are sometimes implanted with subdural electrocorticography (eCoG) recording grids that are used to locate the epileptic focus for subsequent surgical removal. These grids often cover lateral temporal cortex including STS since surgeons wish to avoid removal of the nearby language cortex (for example coverage, see e.g., Tsuchiya, Kawasaki, Oya, Howard & Adolphs, 2008; Flinker, Chang, Barbaro, Berger & Knight, 2011). Although eCoG cannot be used to isolate the spiking of single neurons, the method can be used to measure multiunit activity (MUA) and local field potentials (LFP), two measures that are also available in macaque electrophysiology. Unfortunately, no macaque studies have reported MUA or LFP responses to social cues. It would be valuable to first obtain this macaque MUA and LFP data from macaque anterior STS and to then compare these effects to human eCoG data. Note that the macaque MUA and LFP responses would be available through re-analysis of existing macaque electrophysiology recordings, since MUA, LFP and spiking all obtain from different filtering operations applied to the same electrophysiological recordings. Measures from LFP would be particularly valuable since this electrophysiological signal is most closely related with BOLD fMRI responses in systematic comparisons (Logothetis, Pauls, Augath, Trinath & Oeltermann, 2001; Goense & Logothetis, 2008). The combined analysis of macaque and human LFP and fMRI responses to the same social cues would provide an unprecedented look at both function and homology of social cue representations in STS.

6.4.5 Toward computational theories of gaze discrimination

There is a wealth of behavioural data on how humans discriminate gaze direction and other social cues. Similarly, the fMRI experiments in this thesis join emerging efforts to better characterise neural codes for the direction of social cues. As the empirical findings accumulate, there is a great need for computational models. The field is currently dependent on conceptual models such as the scleral ratio model of gaze di-

rection discrimination (Ando, 2002, see Section 1.2.5). Similarly, electrophysiological data has only been explained by a general framework of hierarchical inhibition across gaze, head and body posture cells (Perrett & Emery, 1994). Such conceptual models may account well for qualitative effects such as the integration of different social cues into a single direction of social attention (Perrett & Emery, 1994). However, the field also contains reports of quantitative effects such as the biasing effects of head view on gaze direction discrimination (Gibson & Pick, 1963) or the widening of gaze directions judged as conveying eye contact in the context of hearing one's own as opposed to another's name (Stoyanova et al., 2010). Such differences in degrees rather than in kind may be better suited to the continuous predictions of a computational account.

Outside of psychology and neuroscience, researchers in computer vision have begun to develop computational models for estimating gaze direction. Here I describe two of these models in order to provide some insight into what a psychological or neuroscientific computational model of social cue representations might look like.

Conventional eye trackers analyse data from high-speed infra-red cameras. In general, such gaze analysis involves first ensuring that the relationship between the camera and the eye is fixed and then learning a view-based mapping between pupil position in the captured image and target eccentricity on the participant's screen (Weidenbacher et al., 2006). This analysis reveals little about how humans perceive gaze in the visible colour spectrum, without a fixed relationship between actor and observer. However, recent computer vision models have become more relevant to human performance. These models get closer to the problems faced by real human performance by estimating gaze direction with conventional web cameras that operate in the visible colour spectrum, without a fixed relationship between the camera and the actor's eyes (Weidenbacher et al., 2006; Morency et al., 2006).

Weidenbacher et al. (2006) developed a two-stage model for carrying out gaze discrimination based on video data. The model first estimates head view and after this, eye region information is estimated by fitting a set of gabor filters to the eye region (Figure 6.1). Intuitively, such gabors represent a view-based eye region model where the position of the iris in the sclera is given by the phase of the gabor. Gaze direction is mapped through a look-up table where each combination of eye region gabor phase and estimated head view gives a certain gaze estimate. The model achieved reasonable gaze discrimination performance, with gaze angle errors of 1-10° (190 cm viewing

distance). Mirroring human performance, accuracy was reduced for far eccentric gaze directions (cf. Section 1.2.1) and for gaze discrimination in turned relative to front-facing heads (cf. Section 1.2.3). However, the model does not predict repulsive or attractive effects of head view on gaze estimates (cf. Gibson & Pick, 1963; Anstis et al., 1969; Cline, 1967). Note also that the proposed gabor filter approach would have to be extended to account for vertically averted gaze.

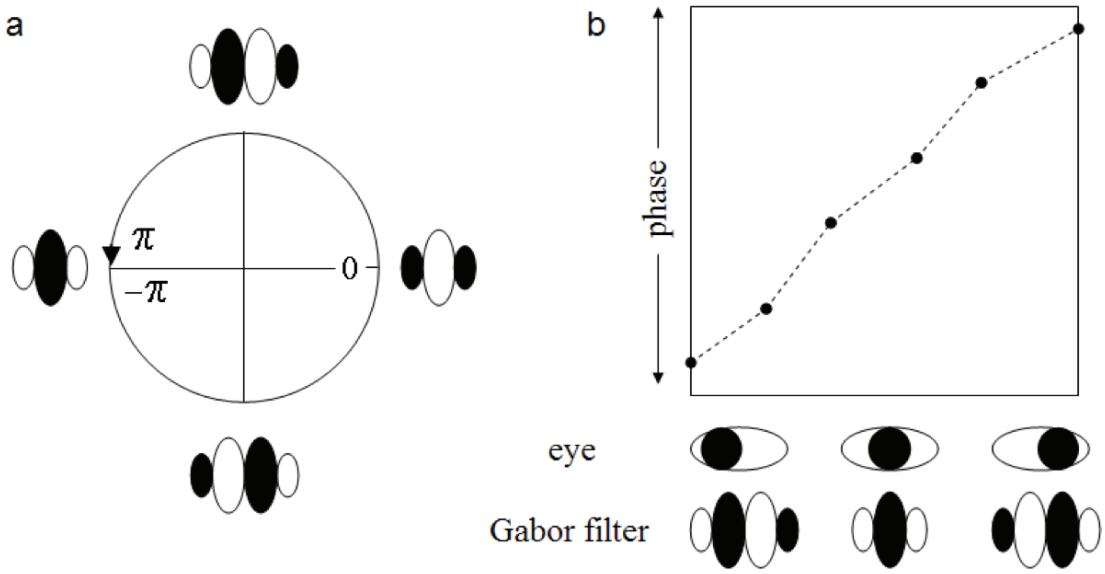


Figure 6.1: Illustration of a computational representation of gaze direction based on the phase of a gabor. (**a**) The gabor filter is generated by multiplying a wave function with a Gaussian envelope. The filter shape changes with the phase of the wave function. White represents positive filter components and black represents negative components. (**b**) The gabor filter is fitted to the eye image. The phase of the best-fitting gabor filter is an estimate of gaze direction. Adapted from Weidenbacher et al. (2006).

A related model by Morency et al. (2006) used a machine learning approach where the actor's left and right eye regions were isolated based on head view estimation and a pattern classifier was trained to recognise gaze direction based on the image pixels inside the segmented eye regions. Importantly, the pattern classifier had no information about the actor's head view. Gaze discrimination performance for this model appears comparable to the two-stage model by Weidenbacher et al. (2006), but unlike that model, vertically averted gaze was also accommodated. However, performance was only tested in the context of videos with very slight head view changes. It remains

to be determined whether this model would perform reasonably also in the context of large head view changes, where conceptual view-based eye region models such as the scleral ratio model fail.

These computer vision models demonstrate that generative computational models for social cue perception are possible with existing technology. However, computer vision researchers create models with the aim of achieving complete accuracy, while the gaze perception phenomena that psychologists and neuroscientists wish to explain often reflect interesting failures in gaze perception performance such as the biasing effects of head view on gaze direction discrimination (Gibson & Pick, 1963). Thus, the models from computer vision can inspire but not replace psychological and neuroscientific models of human gaze direction representations. The development of such models would be particularly timely since the continuous predictions that computational models generate can readily form predictor RDMs for RSA.

6.5 Conclusions

In this thesis I used MVPA methods combined with fMRI scans and behavioural testing to study how humans and macaques discriminate the direction of another person's attention as conveyed by head turn, gaze direction and head view. Across all experiments, I identified anterior STS as a key neural correlate of such directional social cue representations. Responses in this region were also found to discriminate gaze direction across different head views, suggesting that human anterior STS represents the direction of another's attention regardless of how this is conveyed. Macaque fMRI experiments provided further support for the role of anterior STS in discriminating gaze direction and head view, thus providing an indication that anterior STS representations of social cues are similar across these species when similar data acquisition and analysis methods are used. This research contributes to an enhanced understanding of how MVPA can be used to study high-level, socially-relevant perceptual processes and, specifically, how STS represents the direction of another's attention.

References

Addams, R. (1834). An account of a peculiar optical phenomenon seen after having looked at a moving body. *London Edinburgh Philosophical Magazine and Journal of Science*, *5*, 373–374.

Allison, T., Puce, A., & Mccarthy, G. (2000). Social perception from visual cues: Role of the STS region. *Trends in Cognitive Sciences*, *4*(7), 267–278.

Ando, S. (2002). Luminance-induced shift in the apparent direction of gaze. *Perception*, *31*(6), 657–674.

Ando, S. (2004). Perception of gaze direction based on luminance ratio. *Perception*, *33*(10), 1173–1184.

Andrews, T. J. (2005). Visual cortex: How are faces and objects represented? *Current Biology*, *15*, R451–R453.

Andrews, T. J. & Ewbank, M. P. (2004). Distinct representations for facial identity and changeable aspects of faces in the human temporal lobe. *NeuroImage*, *23*, 905–913.

Anstis, S. M., Mayhew, J. W., & Morley, T. (1969). The perception of where a face or television 'portrait' is looking. *American Journal of Psychology*, *82*(4), 474–489.

Arcaro, M., Mcmains, S., Singer, B., & Kastner, S. (2009). Retinotopic organization of human ventral visual cortex. *Journal of Neuroscience*, *29*(34), 10638.

Argyle, M., Lefebvre, L., & Cook, M. (1974). The meaning of five patterns of gaze. *European Journal of Social Psychology*, *4*(2), 125–136.

Bakola, S., Gregoriou, G., Moschovakis, A., Raos, V., & Savaki, H. (2007). Saccade-related information in the superior temporal motion complex: Quantitative functional mapping in the monkey. *Journal of Neuroscience*, *27*(9), 2224–2229.

Baldwin, D. A. (1991). Infants's contribution to the achievement of joint reference. *Child Development*, *62*(5), 875–890.

Baldwin, D. A. (1993a). Early referential understanding: Infants' ability to recognize referential acts for what they are. *Developmental Psychology*, *29*(5), 832–843.

Baldwin, D. A. (1993b). Infants' ability to consult the speaker for clues to word reference. *Journal of Child Language*, *20*(2), 395–418.

Barlow, H. B. & Hill, R. M. (1963). Evidence for a physiological explanation of the waterfall phenomenon and figural after-effects. *Nature*, *200*(491), 1345–1347.

Bartels, A., Logothetis, N. K., & Moutoussis, K. (2008). fMRI and its interpretations: An illustration on directional selectivity in area V5/MT. *Trends in Neurosciences*, *31*(9), 444–453.

Baylis, G., Rolls, E., & Leonard, C. (1985). Selectivity between faces in the responses of a population of neurons in the cortex in the superior temporal sulcus of the monkey. *Brain Research*, *342*, 91–102.

Bi, T., Su, J., Chen, J., & Fang, F. (2009). The role of gaze direction in face viewpoint aftereffect. *Vision Research*, *49*, 2322–2327.

Bodurka, J., Ye, F., Petridou, N., Murphy, K., & Bandettini, P. A. (2007). Mapping the MRI voxel volume in which thermal noise matches physiological noise - implications for fMRI. *NeuroImage*, *34*, 542–549.

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

Brennan, S. E., Chen, X., Dickinson, C. a., Neider, M. B., & Zelinsky, G. J. (2008). Coordinating cognition: The costs and benefits of shared gaze during collaborative search. *Cognition*, *106*(3), 1465–77.

Brewer, A., Liu, J., Wade, A., & Wandell, B. (2005). Visual field maps and stimulus selectivity in human ventral occipital cortex. *Nature Neuroscience*, *8*, 1102–1109.

Brothers, L. & Ring, B. (1993). Mesial temporal neurons in the macaque monkey with responses selective for aspects of social stimuli. *Behavioural Brain Research*, *57*(1), 53–61.

Bruce, C., Desimone, R., & Gross, C. (1981). Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *Journal of Neurophysiology*, *46*, 369–384.

Bruce, V. & Young, A. W. (1986). Understanding face recognition. *British Journal of Psychology*, *77 (Pt 3)*, 305–27.

Burgoon, J. K., Manusov, V., Mineo, P., & Hale, J. L. (1985). Effects of gaze on hiring, credibility, attraction and relational message interpretation. *Journal of Nonverbal Behavior*, *9*(3), 133–146.

Caggiano, V., Fogassi, L., Rizzolatti, G., Pomper, J. K., Thier, P., Giese, M. A., & Casile, A. (2011). View-based encoding of actions in mirror neurons of area F5 in macaque premotor cortex. *Current Biology*, *21*(2), 144–8.

Calder, A. J., Beaver, J., Winston, J. S., Dolan, R. J., Jenkins, R., Eger, E., & Henson, R. N. A. (2007). Separate coding of different gaze directions in the superior temporal sulcus and inferior parietal lobule. *Current Biology*, *17*, 20–25.

Calder, A. J., Jenkins, R., Cassel, A., & Clifford, C. (2008). Visual representation of eye gaze is coded by a nonopponent multichannel system. *Journal of Experimental Psychology: General*, *137*, 244–261.

Calder, A. J., Lawrence, A., Keane, J., Scott, S., Owen, A., Christoffels, I., & Young, A. W. (2002). Reading the mind from eye gaze. *Neuropsychologia*, *40*, 1129–1138.

Campbell, R., Heywood, C., Cowey, A., Regard, M., & Landis, T. (1990). Sensitivity to eye gaze in prosopagnosic patients and monkeys with superior temporal sulcus ablation. *Neuropsychologia*, *28*, 1123–1142.

Carlin, J. D., Calder, A. J., Kriegeskorte, N., Nili, H., & Rowe, J. B. (2011). A head view-invariant representation of gaze direction in anterior superior temporal sulcus. *Current Biology*, *21*(21), 1817–1821.

Carlin, J. D., Rowe, J. B., Kriegeskorte, N., Thompson, R., & Calder, A. J. (2011). Direction-sensitive codes for observed head turns in human superior temporal sulcus. *Cerebral Cortex*.

Cavanna, A. E. & Trimble, M. R. (2006). The precuneus: A review of its functional anatomy and behavioural correlates. *Brain*, *129*, 564–83.

Chan, A. W.-Y. & Downing, P. E. (2011). Faces and eyes in human lateral prefrontal cortex. *Frontiers in Human Neuroscience*, *5*, 1–10.

Chen, X., Pereira, F., Lee, W., Strother, S., & Mitchell, T. (2006). Exploring predictive and reproducible modeling with the single-subject FIAC dataset. *Human Brain Mapping*, *27*, 452–461.

Chumbley, J. & Friston, K. J. (2009). False discovery rate revisited: FDR and topological inference using Gaussian random fields. *NeuroImage*, *44*(1), 62–70.

Clifford, C. (2002). Perceptual adaptation: Motion parallels orientation. *Trends in Cognitive Sciences*, *6*, 136–143.

Cline, M. (1967). The perception of where a person is looking. *American Journal of Psychology*, *80*(1), 41–50.

Conty, L., Tijus, C., Hugueville, L., Coelho, E., & George, N. (2006). Searching for asymmetries in the detection of gaze contact versus averted gaze under different head views: A behavioural study. *Spatial Vision*, *19*(6), 529–45.

Cusack, R., Brett, M., & Osswald, K. (2003). An evaluation of the use of magnetic field maps to undistort echo-planar images. *NeuroImage*, *18*, 127–142.

Damasio, H., Tranel, D., Grabowski, T., Adolphs, R., & Damasio, a. (2004). Neural systems behind word and concept retrieval. *Cognition*, *92*(1-2), 179–229.

Davidenko, N., Remus, D. a., & Grill-Spector, K. (2011). Face-likeness and image variability drive responses in human face-selective ventral regions. *Human Brain Mapping*.

De Souza, W., Eifuku, S., Tamura, R., Nishijo, H., & Ono, T. (2005). Differential characteristics of face neuron responses within the anterior superior temporal sulcus of macaques. *Journal of Neurophysiology*, *94*, 1252–1266.

Deaner, R. & Platt, M. (2003). Reflexive social attention in monkeys and humans. *Current Biology*, *13*, 1609–1613.

Decety, J. & Lamm, C. (2007). The role of the right temporoparietal junction in social interaction: How low-level computational processes contribute to meta-cognition. *Neuroscientist*, *13*, 580–593.

Dimigen, O., Valsecchi, M., Sommer, W., & Kliegl, R. (2009). Human microsaccade-related visual brain responses. *Journal of Neuroscience*, *29*(39), 12321–12331.

Dodgson, N. A. (2004). Variation and extrema of human interpupillary distance. *Stereoscopic Displays and Applications XI*, *5291*, 36–46.

Downing, P. E., Dodds, C., & Bray, D. (2004). Why does the gaze of others direct visual attention? *Visual Cognition*, *11*, 71–79.

Driver, J., Davis, G., Kidd, P., Maxwell, E., & Baron-Cohen, S. (1999). Gaze perception triggers reflexive visuospatial orienting. *Visual Cognition*, *6*(5), 509–540.

Dumoulin, S. O., Bittar, R. G., Kabani, N. J., Baker, C. L., Le Goualher, G., Pike, G. B., & Evans, A. C. (2000). A new anatomical landmark for reliable identification of human area V5/MT: A quantitative analysis of sulcal patterning. *Cerebral Cortex*, *10*(5), 454–63.

Eacott, M., Heywood, C., Gross, C., & Cowey, A. (1993). Visual discrimination impairments following lesions of the superior temporal sulcus are not specific for facial stimuli. *Neuropsychologia*, *31*, 609–619.

Ecker, C., Marquand, A., Mourao-Miranda, J., Johnston, P., Daly, E. M., Brammer, M. J., Maltezos, S., Murphy, C. M., Robertson, D., Williams, S. C., & Murphy,

D. G. M. (2010). Describing the brain in autism in five dimensions - Magnetic resonance imaging-assisted diagnosis of autism spectrum disorder using a multiparameter classification approach. *Journal of Neuroscience*, *30*(32), 10612–10623.

Emery, N. J. (2000). The eyes have it: The neuroethology, function and evolution of social gaze. *Neuroscience and Biobehavioral Reviews*, *24*(6), 581–604.

Engbert, R. & Kliegl, R. (2003). Microsaccades uncover the orientation of covert attention. *Vision Research*, *43*, 1035–1045.

Engell, A. & Haxby, J. V. (2007). Facial expression and gaze-direction in human superior temporal sulcus. *Neuropsychologia*, *45*, 3234–3241.

Evans, A., Zilles, K., Lancaster, J., Martinez, M., Mazziotta, J., Fox, P., Tordesillas-Gutierrez, D., & Salinas, F. (2007). Bias between MNI and Talairach coordinates analyzed using the ICBM-152 brain template. *Human Brain Mapping*, *28*(11), 1194–1205.

Ewbank, M. & Andrews, T. J. (2008). Differential sensitivity for viewpoint between familiar and unfamiliar faces in human visual cortex. *NeuroImage*, *40*, 1857–1870.

Ewbank, M. P., Jennings, C., & Calder, A. J. (2009). Why are you angry with me? Facial expressions of threat influence perception of gaze direction. *Journal of Vision*, *9*(12), 1–7.

Fang, F. & He, S. (2005). Viewer-centered object representation in the human visual system revealed by viewpoint aftereffects. *Neuron*, *45*(5), 793–800.

Fang, F. & Ijichi, K. (2007). Transfer of the face viewpoint aftereffect from adaptation to different and inverted faces. *Journal of Vision*, *7*, 1–9.

Ferrari, P., Kohler, E., Fogassi, L., & Gallese, V. (2000). The ability to follow eye gaze and its emergence during development in macaque monkeys. *Proceedings of the National Academy of Sciences*, *97*, 13997–14002.

Fischl, B., Sereno, M. I., & Tootell, R. (1999). High-resolution intersubject averaging and a coordinate system for the cortical surface. *Human Brain Mapping*, *284*, 272–284.

Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Segonne, F., Salat, D. H., Busa, E., Seidman, L. J., Goldstein, J., Kennedy, D., Caviness, V., Makris, N., Rosen, B., & Dale, A. M. (2004). Automatically parcellating the human cerebral cortex. *Cerebral Cortex*, *14*(1), 11–22.

Flinker, A., Chang, E. F., Barbaro, N. M., Berger, M. S., & Knight, R. T. (2011). Sub-centimeter language organization in the human temporal lobe. *Brain and Language*, *117*(3), 103–9.

Fox, C., Iaria, G., & Barton, J. (2009). Defining the face processing network: Optimization of the functional localizer in fMRI. *Human Brain Mapping*, *30*, 1637–1651.

Freeman, J., Brouwer, G. J., Heeger, D. J., & Merriam, E. P. (2011). Orientation decoding depends on maps, not columns. *Journal of Neuroscience*, *31*(13), 4792–4804.

Freiwald, W. A. & Tsao, D. Y. (2010). Functional compartmentalization and viewpoint generalization within the macaque face-processing system. *Science*, *330*(6005), 845–851.

Friesen, C. & Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin & Review*, *5*, 490–495.

Frischen, A., Bayliss, A., & Tipper, S. P. (2007). Gaze cueing of attention: Visual attention, social cognition, and individual differences. *Psychological Bulletin*, *133*, 694–724.

Friston, K. J., Rotshtein, P., Geng, J., Sterzer, P., & Henson, R. N. A. (2006). A critique of functional localisers. *NeuroImage*, *30*, 1077–1087.

Frith, C. D. & Frith, U. (2008). Implicit and explicit processes in social cognition. *Neuron*, *60*, 503–510.

Furmanski, C. S. & Engel, S. a. (2000). An oblique effect in human primary visual cortex. *Nature Neuroscience*, *3*(6), 535–6.

Gamer, M. & Hecht, H. (2007). Are you looking at me? Measuring the cone of gaze. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(3), 705–15.

Gardner, J. L. (2010). Is cortical vasculature functionally organized? *NeuroImage*, *49*(3), 1953–6.

Genovese, C. R., Lazar, N. A., & Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage*, *15*(4), 870–878.

George, M., Johnson, K., Lu, X., Jin, B., Strasburger, A., Laken, S., & Kozel, F. (2009). Feature selection for fMRI-based deception detection. *BMC Bioinformatics*, *10*(Suppl 9), S15.

George, N., Driver, J., & Dolan, R. J. (2001). Seen gaze-direction modulates fusiform activity and its coupling with other brain areas during face processing. *NeuroImage*, *13*, 1102–1112.

Gibson, J. & Pick, A. (1963). Perception of another person's looking behavior. *American Journal of Psychology*, *76*(3), 386–394.

Gobbini, M., Koralek, A., Bryan, R., Montgomery, K., & Haxby, J. V. (2007). Two takes on the social brain: A comparison of theory of mind tasks. *Journal of Cognitive Neuroscience*, *19*, 1803–1814.

Goense, J. & Logothetis, N. (2008). Neurophysiology of the BOLD fMRI signal in awake monkeys. *Current Biology*, *18*, 631–640.

Golland, P., Liang, F., Mukherjee, S., & Panchenko, D. (2005). Permutation tests for classification. *Lecture Notes in Computer Science*, *3559*, 36–39.

Gottlieb, J. (2007). From thought to action: The parietal cortex as a bridge between perception, action, and cognition. *Neuron*, *53*(1), 9–16.

Grill-Spector, K., Henson, R. N. A., & Martin, A. (2006). Repetition and the brain: Neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, *10*, 14–23.

Grill-Spector, K., Knouf, N., & Kanwisher, N. G. (2004). The fusiform face area subserves face perception, not generic within-category identification. *Nature Neuroscience*, *7*(5), 555–62.

Grill-Spector, K. & Malach, R. (2001). FMR-adaptation: A tool for studying the functional properties of human cortical neurons. *Acta Psychologica*, *107*, 293–321.

Grosbras, M.-H., Laird, A., & Paus, T. (2005). Cortical regions involved in eye movements, shifts of attention, and gaze perception. *Human Brain Mapping*, *25*, 140–154.

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.

Hanke, M., Halchenko, Y., Sederberg, P., Olivetti, E., Fründ, I., Rieger, J., Herrmann, C., Haxby, J. V., Hanson, S. J., & Pollmann, S. (2009). PyMVPA: A unifying approach to the analysis of neuroscientific data. *Frontiers in Neuroinformatics*, *3*, 1–13.

Hanna, J. & Brennan, S. (2007). Speakers' eye gaze disambiguates referring expressions early during face-to-face conversation. *Journal of Memory and Language*, *57*(4), 596–615.

Hanson, S. J. & Halchenko, Y. (2008). Brain reading using full brain support vector machines for object recognition: There is no "face" identification area. *Neural Computation*, *20*, 486–503.

Hanson, S. J., Matsuka, T., & Haxby, J. V. (2004). Combinatorial codes in ventral temporal lobe for object recognition: Haxby (2001) revisited: Is there a "face" area? *NeuroImage*, *23*(1), 156–166.

Harries, M. & Perrett, D. (1991). Visual processing of faces in temporal cortex: Physiological evidence for a modular organization and possible anatomical correlates. *Journal of Cognitive Neuroscience*, *3*, 10–24.

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.

Hasselmo, M., Rolls, E., Baylis, G., & Nalwa, V. (1989). Object-centered encoding by face-selective neurons in the cortex in the superior temporal sulcus of the monkey. *Experimental Brain Research*, *75*, 417–429.

Haxby, J. V., Gobbini, M., Furey, M., Ishai, A., Schouten, J., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, *293*, 2425–2430.

Haxby, J. V., Hoffman, E., & Gobbini, M. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*, 223–233.

Haynes, J. & Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience*, *7*, 523–534.

Haynes, J.-D. & Rees, G. (2005). Predicting the stream of consciousness from activity in human visual cortex. *Current Biology*, *15*(14), 1301–7.

Hein, G. & Knight, R. (2008). Superior temporal sulcus - it's my area: Or is it? *Journal of Cognitive Neuroscience*, *20*, 2125–2136.

Henson, R. N. A. (2003). Analysis of fMRI timeseries: Linear time-invariant models, event-related fMRI and optimal experimental design. In Frackowiak, R. S. J., Friston, K. J., Frith, C. D., Dolan, R. J., & Price, C. J. (Eds.), *Human Brain Function*, (pp. 793–822)., New York. Academic Press.

Heywood, C. A. & Cowey, A. (1992). The role of the 'face-cell' area in the discrimination and recognition of faces by monkeys. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *335*(1273), 31–7.

Hoffman, E. & Haxby, J. V. (2000). Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nature Neuroscience*, *3*, 80–84.

Hoffman, K., Gothard, K., Schmid, M., & Logothetis, N. (2007). Facial-expression and gaze-selective responses in the monkey amygdala. *Current Biology*, *17*, 766–772.

Holmes, A. & Friston, K. J. (1998). Generalisability, random effects and population inference. *Neuroimage*, *7*, S754.

Huk, A. C., Dougherty, R. F., & Heeger, D. J. (2002). Retinotopy and functional subdivision of human areas MT and MST. *Journal of Neuroscience*, *22*(16), 7195–205.

Jellema, T. & Perrett, D. (2003). Perceptual history influences neural responses to face and body postures. *Journal of Cognitive Neuroscience*, *15*, 961–971.

Jenkins, R. (2007). The lighter side of gaze perception. *Perception*, *36*(8), 1266–1268.

Jenkins, R., Beaver, J., & Calder, A. J. (2006). I thought you were looking at me: Direction-specific aftereffects in gaze perception. *Psychological Science*, *17*, 506–513.

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. M. (2002). Improved optimisation for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, *17*, 825–841.

Kamitani, Y. & Sawahata, Y. (2010). Spatial smoothing hurts localization but not information: Pitfalls for brain mappers. *NeuroImage*, *49*(3), 1949–1952.

Kamitani, Y. & Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nature Neuroscience*, *8*, 679–685.

Kamitani, Y. & Tong, F. (2006). Decoding seen and attended motion directions from activity in the human visual cortex. *Current Biology*, *16*, 1096–1102.

Kamphuis, S., Dicke, P. W., & Thier, P. (2009). Neuronal substrates of gaze following in monkeys. *European Journal of Neuroscience*, *29*, 1732–1738.

Kanwisher, N. G. (2000). Domain specificity in face perception. *Nature Neuroscience*, *3*(8), 759–63.

Kanwisher, N. G. (2010). Functional specificity in the human brain: A window into the functional architecture of the mind. *Proceedings of the National Academy of Sciences*.

Kanwisher, N. G., McDermott, J., & Chun, M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, *17*, 4302–4311.

Kanwisher, N. G. & Yovel, G. (2006). The fusiform face area: A cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *361*, 2109–2128.

Kastner, S., Pinsk, M., Weerd, P., Desimone, R., & Ungerleider, L. G. (1999). Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron*, *22*, 751–761.

Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta Psychologica*, *26*(1), 22–63.

Klein, J., Shepherd, S. V., & Platt, M. (2009). Social attention and the brain. *Current Biology*, *19*(20), R958–R962.

Klein, R. (2000). Inhibition of return. *Trends in Cognitive Sciences*, *4*(4), 138–147.

Kluttz, N., Mayes, B., West, R., & Kerby, D. (2009). The effect of head turn on the perception of gaze. *Vision Research*, *49*, 1979–1993.

Kobayashi, H. & Kohshima, S. (1997). Unique morphology of the human eye. *Nature*, *387*, 767.

Kravitz, D., Kriegeskorte, N., & Baker, C. I. (2010). High-level visual object representations are constrained by position. *Cerebral Cortex*, *20*, 2916–2925.

Krekelberg, B., Boynton, G., & Vanwezel, R. (2006). Adaptation: From single cells to BOLD signals. *Trends in Neurosciences*, *29*, 250–256.

Kriegeskorte, N. & Bandettini, P. A. (2007). Analyzing for information, not activation, to exploit high-resolution fMRI. *NeuroImage*, *38*, 649–662.

Kriegeskorte, N., Bandettini, P. A., & Cusack, R. (2009). How does an fMRI voxel sample the neuronal activity pattern: Compact-kernel or complex-spatiotemporal filter? *NeuroImage*, *49*(3), 1965–1976.

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*, *104*, 20600–20605.

Kriegeskorte, N., Goebel, R., & Bandettini, P. A. (2006). Information-based functional brain mapping. *Proceedings of the National Academy of Sciences*, *103*, 3863–3868.

Kriegeskorte, N., Mur, M., & Bandettini, P. A. (2008). Representational similarity analysis - connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, *2*, 1–28.

Kriegeskorte, N., Mur, M., Ruff, D., Kiani, R., Bodurka, J., Esteky, H., Tanaka, K., & Bandettini, P. A. (2008). Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron*, *60*, 1126–1141.

Laconte, S., Strother, S., Cherkassky, V., Anderson, J., & Hu, X. (2005). Support vector machines for temporal classification of block design fMRI data. *NeuroImage*, *26*, 317–329.

Langton, S., Watt, R., & Bruce, I. (2000). Do the eyes have it? Cues to the direction of social attention. *Trends in Cognitive Sciences*, *4*, 50–59.

Langton, S. R. & Bruce, V. (2000). You must see the point: Automatic processing of cues to the direction of social attention. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(2), 747–57.

Langton, S. R. H., Honeyman, H., & Tessler, E. (2004). The influence of head contour and nose angle on the perception of eye-gaze direction. *Perception & Psychophysics*, *66*(5), 752–771.

Larsson, J. & Heeger, D. (2006). Two retinotopic visual areas in human lateral occipital cortex. *Journal of Neuroscience*, *26*, 13128–13142.

Laube, I., Kamphuis, S., Dicke, P. W., & Thier, P. (2011). Cortical processing of head- and eye-gaze cues guiding joint social attention. *NeuroImage*, *54*, 1643–1653.

Lawson, R., Clifford, C. W. G., & Calder, A. J. (2009). About turn: The visual representation of human body orientation revealed by adaptation. *Psychological Science*, *20*, 363–371.

Lawson, R. P., Clifford, C. W. G., & Calder, A. J. (2011). A real head turner: Horizontal and vertical head directions are multichannel coded. *Journal of Vision*, *11*, 1–17.

Lee, L. C., Andrews, T. J., Johnson, S. J., Woods, W., Gouws, A., Green, G. G. R., & Young, A. W. (2010). Neural responses to rigidly moving faces displaying shifts in social attention investigated with fMRI and MEG. *Neuropsychologia*, *48*(2), 477–490.

Leite, F. P., Tsao, D. Y., Vanduffel, W., Fize, D., Sasaki, Y., Wald, L. L., Dale, A. M., Kwong, K. K., Orban, G. A., Rosen, B. R., Tootell, R. B. H., & Mandeville, J. B. (2002). Repeated fMRI using iron oxide contrast agent in awake, behaving macaques at 3 Tesla. *NeuroImage*, *16*(2), 283–94.

Lemieux, L., Salekhaddadi, A., Lund, T., Laufs, H., & Carmichael, D. (2007). Modelling large motion events in fMRI studies of patients with epilepsy. *Magnetic Resonance Imaging*, *25*, 894–901.

Leopold, D., O'Toole, A., Vetter, T., & Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level aftereffects. *Nature Neuroscience*, *4*, 89–94.

Lobmaier, J., Tiddeman, B., & Perrett, D. (2008). Emotional expression modulates perceived gaze direction. *Emotion*, *8*, 573–577.

Lobmaier, J. S., Fischer, M. H., & Schwaninger, A. (2006). Objects capture perceived gaze direction. *Experimental Psychology*, *53*(2), 117–122.

Logothetis, N. (2008). What we can do and what we cannot do with fMRI. *Nature*, *453*, 869–878.

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, a. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature*, *412*(6843), 150–7.

Lord, C. & Haith, M. M. (1974). The perception of eye contact. *Perception & Psychophysics*, *16*(3), 413–416.

Mannion, D., McDonald, J., & Clifford, C. (2009). Discrimination of the local orientation structure of spiral Glass patterns early in human visual cortex. *NeuroImage*, *46*(2), 511–515.

Mansfield, E. P., Farroni, T., & Johnson, M. H. (2003). Does gaze perception facilitate overt orienting? *Visual Cognition*, *10*(1), 7–14.

Mikami, A., Newsome, W., & Wurtz, R. (1986a). Motion selectivity in macaque visual cortex. I. Mechanisms of direction and speed selectivity in extrastriate area MT. *Journal of Neurophysiology*, *55*, 1308–1327.

Mikami, A., Newsome, W., & Wurtz, R. (1986b). Motion selectivity in macaque visual cortex. II. Spatiotemporal range of directional interactions in MT and V1. *Journal of Neurophysiology*, *55*, 1328–1339.

Misaki, M., Kim, Y., Bandettini, P. A., & Kriegeskorte, N. (2010). Comparison of multivariate classifiers and response normalizations for pattern-information fMRI. *NeuroImage*, *53*, 103–118.

Mitchell, T., Hutchinson, R., Niculescu, R., Pereira, F., & Wang, X. (2004). Learning to decode cognitive states from brain images. *Machine Learning*, *57*, 145–175.

Mitchell, T. M. (1997). *Machine Learning*. New York: McGraw-Hill.

Miyawaki, Y., Uchida, H., Yamashita, O., Sato, M., Morito, Y., Tanabe, H., Sadato, N., & Kamitani, Y. (2008). Visual image reconstruction from human brain activity using a combination of multiscale local image decoders. *Neuron*, *60*, 915–929.

Morency, L.-P., Christoudias, C. M., & Darrell, T. (2006). Recognizing gaze aversion gestures in embodied conversational discourse. *Proceedings of the 8th international conference on Multimodal interfaces - ICMI '06*, 287.

Mourao-Miranda, J., Bokde, A., Born, C., Hampel, H., & Stetter, M. (2005). Classifying brain states and determining the discriminating activation patterns: Support vector machine on functional MRI data. *NeuroImage*, *28*, 980–995.

Mumford, J. A., Turner, B. O., Ashby, F. G., & Poldrack, R. A. (2011). Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. *NeuroImage*.

Mur, M., Bandettini, P. A., & Kriegeskorte, N. (2009). Revealing representational content with pattern-information fMRI - an introductory guide. *Social Cognitive and Affective Neuroscience*, *4*, 101–109.

Natu, V. S., Jiang, F., Narvekar, A., Keshvari, S., Blanz, V., & O'Toole, A. J. (2010). Dissociable neural patterns of facial identity across changes in viewpoint. *Journal of Cognitive Neuroscience*, *22*(7), 1570–1582.

Nelissen, K., Vanduffel, W., & Orban, G. (2006). Charting the lower superior temporal region, a new motion-sensitive region in monkey superior temporal sulcus. *Journal of Neuroscience*, *26*, 5929–5947.

Nichols, T. & Hayasaka, S. (2003). Controlling the familywise error rate in functional neuroimaging: A comparative review. *Statistical Methods in Medical Research*, *12*(5), 419–46.

Nichols, T. & Holmes, A. (2001). Nonparametric permutation tests for functional neuroimaging: A primer with examples. *Human Brain Mapping*, *15*, 1–25.

Noble, W. S. (2006). What is a support vector machine? *Nature Biotechnology*, *24*(12), 1565–7.

Nummenmaa, L. & Calder, A. J. (2009). Neural mechanisms of social attention. *Trends in Cognitive Sciences*, *13*(3), 135–143.

Oosterhof, N. N., Wiestler, T., Downing, P. E., & Diedrichsen, J. (2010). A comparison of volume-based and surface-based multi-voxel pattern analysis. *NeuroImage*, *56*, 593–600.

Op de Beeck, H. P. (2010). Against hyperacuity in brain reading: Spatial smoothing does not hurt multivariate fMRI analyses? *NeuroImage*, *49*, 1943–1948.

Orban, G. A., Van Essen, D., & Vanduffel, W. (2004). Comparative mapping of higher visual areas in monkeys and humans. *Trends in Cognitive Sciences*, *8*(7), 315–24.

Pageler, N. M., Menon, V., Merin, N. M., Eliez, S., Brown, W. E., & Reiss, A. L. (2003). Effect of head orientation on gaze processing in fusiform gyrus and superior temporal sulcus. *NeuroImage*, *20*(1), 318–329.

Pelphrey, K. A. & Carter, E. (2008). Brain mechanisms for social perception: Lessons from autism and typical development. *Annals of the New York Academy of Sciences*, *1145*, 283–299.

Pelphrey, K. A. & Morris, J. (2006). Brain mechanisms for interpreting the actions of others from biological-motion cues. *Current Directions in Psychological Science*, *15*, 136–140.

Pelphrey, K. A., Singerman, J., Allison, T., & McCarthy, G. (2003). Brain activation evoked by perception of gaze shifts: The influence of context. *Neuropsychologia*, *41*, 156–170.

Pelphrey, K. A., Viola, R., & Mccarthy, G. (2004). When strangers pass: Processing of mutual and averted social gaze in the superior temporal sulcus. *Psychological Science*, *15*, 598–603.

Pereira, F. & Botvinick, M. (2011). Information mapping with pattern classifiers: A comparative study. *NeuroImage*, *56*, 476–496.

Pereira, F., Mitchell, T., & Botvinick, M. (2009). Machine learning classifiers and fMRI: A tutorial overview. *NeuroImage*, *45*, S199–S209.

Perrett, D., Oram, M., Harries, M., Bevan, R., Hietanen, J., Benson, P., & Thomas, S. (1991). Viewer-centred and object-centred coding of heads in the macaque temporal cortex. *Experimental Brain Research*, *86*, 159–173.

Perrett, D., Rolls, E., & Caan, W. (1982). Visual neurones responsive to faces in the monkey temporal cortex. *Experimental Brain Research*, *47*(3), 329–342.

Perrett, D., Smith, P., Potter, D., Mistlin, A., Head, A., Milner, A., & Jeeves, M. (1985). Visual cells in the temporal cortex sensitive to face view and gaze direction. *Proceedings of the Royal Society, B: Biological Sciences*, *223*, 293–317.

Perrett, D., Smith, P. A., Potter, D. D., Mistlin, A. J., Head, A. S., Milner, A. D., & Jeeves, M. A. (1984). Neurones responsive to faces in the temporal cortex: Studies of functional organization, sensitivity to identity and relation to perception. *Human Neurobiology*, *3*(4), 197–208.

Perrett, D. I. & Emery, N. J. (1994). Understanding the intentions of others from visual signals: Neurophysiological evidence. *Cahiers de Psychologie Cognitive*, *13*, 683–694.

Perrett, D. I., Hietanen, J. K., Oram, M. W., & Benson, P. J. (1992). Organization and functions of cells responsive to faces in the temporal cortex. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *335*(1273), 23–30.

Perrett, D. I., Smith, P. A. J., Mistlin, A. J., Chitty, A. J., Head, A. S., Potter, D. D., Broennimann, R., Milner, A. D., & Jeeves, M. A. (1985). Visual analysis of body movements by neurones in the temporal cortex of the macaque monkey: A preliminary report. *Behavioural Brain Research*, *16*(2-3), 153–70.

Pinsk, M., DeSimone, K., Moore, T., Gross, C., & Kastner, S. (2005). Representations of faces and body parts in macaque temporal cortex: A functional MRI study. *Proceedings of the National Academy of Sciences*, *102*, 6996–7001.

Pinsk, M. A., Arcaro, M., Weiner, K. S., Kalkus, J. F., Inati, S. J., Gross, C. G., & Kastner, S. (2009). Neural representations of faces and body parts in macaque and human cortex: A comparative fMRI study. *Journal of Neurophysiology*, *101*, 2581–2600.

Pitcher, D., Dilks, D. D., Saxe, R. R., Triantafyllou, C., & Kanwisher, N. G. (2011). Differential selectivity for dynamic versus static information in face-selective cortical regions. *NeuroImage*, *56*, 2356–2363.

Puce, A., Allison, T., Bentin, S., Gore, J. C., & McCarthy, G. (1998). Temporal cortex activation in humans viewing eye and mouth movements. *Journal of Neuroscience*, *18*(6), 2188–99.

Rajimehr, R., Young, J., & Tootell, R. B. H. (2009). An anterior temporal face patch in human cortex, predicted by macaque maps. *Proceedings of the National Academy of Sciences*, *106*, 1995–2000.

Rhodes, G. & Jeffery, L. (2006). Adaptive norm-based coding of facial identity. *Vision research*, *46*(18), 2977–87.

Rhodes, G., Jeffery, L., Watson, T. L., Clifford, C. W. G., & Nakayama, K. (2003). Fitting the mind to the world: Face adaptation and attractiveness aftereffects. *Psychological Science*, *14*(6), 558–66.

Ricciardelli, P., Baylis, G., & Driver, J. (2000). The positive and negative of human expertise in gaze perception. *Cognition*, *77*(1), B1–14.

Rowe, J. B., Eckstein, D., Braver, T., & Owen, A. (2008). How does reward expectation influence cognition in the human brain? *Journal of Cognitive Neuroscience*, *20*(11), 1980–1992.

Sapountzis, P., Schluppeck, D., Bowtell, R., & Peirce, J. W. (2010). A comparison of fMRI adaptation and multivariate pattern classification analysis in visual cortex. *NeuroImage*, *49*(2), 1632–1640.

Sasaki, Y., Rajimehr, R., Kim, B., Ekstrom, L. B., Vanduffel, W., & Tootell, R. B. H. (2006). The radial bias: A different slant on visual orientation sensitivity in human and nonhuman primates. *Neuron*, *51*, 661–670.

Saxe, R., Xiao, D., Kovacs, G., Perrett, D., & Kanwisher, N. G. (2004). A region of right posterior superior temporal sulcus responds to observed intentional actions. *Neuropsychologia*, *42*, 1435–1446.

Scaife, M. & Bruner, J. (1975). The capacity for joint visual attention in the infant. *Nature*, *253*, 265–266.

Scholz, J., Triantafyllou, C., Whitfield-Gabrieli, S., Brown, E., & Saxe, R. (2009). Distinct regions of right temporo-parietal junction are selective for theory of mind and exogenous attention. *PLoS ONE*, *4*, e4869.

Schwarzlose, R. F., Swisher, J. D., Dang, S., & Kanwisher, N. G. (2008). The distribution of category and location information across object-selective regions in human visual cortex. *Proceedings of the National Academy of Sciences*, *105*(11), 4447–52.

Seltzer, B. & Pandya, D. (1994). Parietal, temporal, and occipital projections to cortex of the superior temporal sulcus in the rhesus monkey: A retrograde tracer study. *Journal of Comparative Neurology*, *343*, 445–463.

Senju, A. & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology*, *18*, 668–671.

Senju, A. & Hasegawa, T. (2005). Direct gaze captures visuospatial attention. *Visual Cognition*, *12*, 127–144.

Senju, A. & Johnson, M. H. (2009). The eye contact effect: mechanisms and development. *Trends Cogn Sci*, *13*(3), 127–134.

Seyama, J. & Nagayama, R. (2006). Eye direction aftereffect. *Psychological Research*, *70*, 59–67.

Seymour, K., Clifford, C., Logothetis, N., & Bartels, A. (2009). The coding of color, motion, and their conjunction in the human visual cortex. *Current Biology*, *19*, 177–183.

Shepherd, S. V. (2010). Following gaze: Gaze-following behavior as a window into social cognition. *Frontiers in Integrative Neuroscience*, *4*, 1–5.

Shepherd, S. V., Klein, J., Deaner, R., & Platt, M. (2009). Mirroring of attention by neurons in macaque parietal cortex. *Proceedings of the National Academy of Sciences*, *106*, 9489–9494.

Shmuel, A., Chaimow, D., Raddatz, G., Ugurbil, K., & Yacoub, E. (2010). Mechanisms underlying decoding at 7 T: Ocular dominance columns, broad structures, and macroscopic blood vessels in V1 convey information on the stimulated eye. *NeuroImage*, *49*(3), 1957–1964.

Sinha, P. (2000). Here's looking at you, kid. *Perception*, *29*(8), 1005–1008.

Sirotin, Y. & Das, A. (2009). Anticipatory haemodynamic signals in sensory cortex not predicted by local neuronal activity. *Nature*, *457*, 475–479.

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., Bannister, P. R., De Luca, M., Drobnjak, I., Flitney, D. E., Niazy, R. K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J. M., & Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, *23*, S208–19.

Snowden, R., Treue, S., & Andersen, R. A. (1992). The response of neurons in areas V1 and MT of the alert rhesus monkey to moving random dot patterns. *Experimental Brain Research*, *88*(2), 389–400.

Spiridon, M. & Kanwisher, N. G. (2002). How distributed is visual category information in human occipito-temporal cortex? An fMRI study. *Neuron*, *35*, 1157–1165.

Stein, T., Senju, A., Peelen, M. V., & Sterzer, P. (2011). Eye contact facilitates awareness of faces during interocular suppression. *Cognition*, *119*(2), 307–311.

Stokes, M., Thompson, R., Cusack, R., & Duncan, J. (2009). Top-down activation of shape-specific population codes in visual cortex during mental imagery. *Journal of Neuroscience*, *29*, 1565–1572.

Stoyanova, R. S., Ewbank, M. P., & Calder, A. J. (2010). "You talkin' to me?": Self-relevant auditory signals influence perception of gaze direction. *Psychological Science*, *21*(12), 1765–1769.

Swisher, J. D., Gatenby, J., Gore, J., Wolfe, B., Moon, C.-H., Kim, S.-G., & Tong, F. (2010). Multiscale pattern analysis of orientation-selective activity in the primary visual cortex. *Journal of Neuroscience*, *30*(1), 325.

Symons, L. A., Lee, K., Cedrone, C. C., & Nishimura, M. (2004). What are you looking at? Acuity for triadic eye gaze. *Journal of General Psychology*, *131*(4), 451–69.

Tazumi, T., Hori, E., Maior, R. S., Ono, T., & Nishijo, H. (2010). Neural correlates to seen gaze-direction and head orientation in the macaque monkey amygdala. *Neuroscience*, *169*, 287–301.

Teufel, C., Alexis, D. M., Todd, H., Lawrance-Owen, A., Clayton, N., & Davis, G. (2009). Social cognition modulates the sensory coding of observed gaze direction. *Current Biology*, *19*, 1274–1277.

Thierry, G., Martin, C., Downing, P. E., & Pegna, A. (2007). Controlling for interstimulus perceptual variance abolishes N170 face selectivity. *Nature Neuroscience*, *10*, 505–511.

Tipples, J. (2002). Eye gaze is not unique: Automatic orienting in response to uninformative arrows. *Psychonomic Bulletin & Review*, *9*, 314–318.

Tipples, J. (2008). Orienting to counterpredictive gaze and arrow cues. *Perception & Psychophysics*, *70*, 77–87.

Todorović, D. (2006). Geometrical basis of perception of gaze direction. *Vision Research*, *46*(21), 3549–62.

Tomasello, M. & Farrar, M. J. (1986). Joint attention and early language. *Child Development*, *57*(6), 1454–63.

Tomasello, M., Hare, B., Lehmann, H., & Call, J. (2007). Reliance on head versus eyes in the gaze following of great apes and human infants: The cooperative eye hypothesis. *Journal of Human Evolution*, *52*, 314–320.

Tranel, D. (2006). Impaired naming of unique landmarks is associated with left temporal polar damage. *Neuropsychology*, *20*(1), 1–10.

Tsao, D. Y., Freiwald, W. A., Knusten, T., Mandeville, J., & Tootell, R. B. H. (2003). Faces and objects in macaque cerebral cortex. *Nature Neuroscience*, *6*, 989–995.

Tsao, D. Y., Freiwald, W. A., Tootell, R. B. H., & Livingstone, M. (2006). A cortical region consisting entirely of face-selective cells. *Science*, *311*, 670–674.

Tsao, D. Y. & Livingstone, M. S. (2008). Mechanisms of face perception. *Annual Review of Neuroscience*, *31*, 411–37.

Tsao, D. Y., Moeller, S., & Freiwald, W. A. (2008). Comparing face patch systems in macaques and humans. *Proceedings of the National Academy of Sciences*, *105*, 19514–19519.

Tsuchiya, N., Kawasaki, H., Oya, H., Howard, M., & Adolphs, R. (2008). Decoding face information in time, frequency and space from direct intracranial recordings of the human brain. *PLoS ONE*, *3*, e3892.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., & Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*, *15*(1), 273–89.

Vander Wyk, B. C., Hudac, C. M., Carter, E. J., Sobel, D. M., & Pelphrey, K. A. (2009). Action understanding in the superior temporal sulcus region. *Psychological Science*, *20*, 771–777.

Vicente-Grabovetsky, A., Mitchell, D. J., Carlin, J. D., & Cusack, R. (2011). Using MVPA to dissociate the role of object-centered and eye-centered reference frames in attention. *Journal of Vision*, *11*(11), 172–172.

Vine, I. (1971). Judgement of direction of gaze: An interpretation of discrepant results. *British Journal of Social and Clinical Psychology*, *10*(4), 320–331.

Von Grüneau, M. & Anston, C. (1995). The detection of gaze direction: A stare-in-the-crowd effect. *Perception*, *24*, 1297–1313.

Wachsmuth, E., Oram, M., & Perrett, D. (1994). Recognition of objects and their component parts: Responses of single units in the temporal cortex of the macaque. *Cerebral Cortex*, *4*, 509–522.

Wang, G., Tanifuji, M., & Tanaka, K. (1998). Functional architecture in monkey inferotemporal cortex revealed by in vivo optical imaging. *Neuroscience Research*, *32*(1), 33–46.

Watanabe, S., Kakigi, R., Miki, K., & Puce, A. (2006). Human MT/V5 activity on viewing eye gaze changes in others: A magnetoencephalographic study. *Brain Research*, *1092*(1), 152–60.

Webster, M., Kaping, D., Mizokami, Y., & Duhamel, P. (2004). Adaptation to natural facial categories. *Nature*, *428*(6982), 557–561.

Webster, M. A. & MacLeod, D. I. A. (2011). Visual adaptation and face perception. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *366*(1571), 1702–1725.

Weidenbacher, U., Layher, G., Bayerl, P., & Neumann, H. (2006). Detection of head pose and gaze direction for human-computer interaction. *Perception and Interactive Technologies*, 9–19.

Wicker, B., Henaff, M.-A., & Decety, J. (1998). Brain regions involved in the perception of gaze : A PET study. *NeuroImage*, *227*(8), 221–227.

Wollaston, W. (1824). On the apparent direction of eyes in a portrait. *Philosophical transactions of the Royal Society of London*, *B114*, 247–256.

Appendix A: Direction-sensitive responses to head turns in human superior temporal sulcus

A reprint of Experiments 1-4 (Chapter 3) in published form.

Cerebral Cortex Advance Access published June 27, 2011

Cerebral Cortex
doi:10.1093/cercor/bhr061

Direction-Sensitive Codes for Observed Head Turns in Human Superior Temporal Sulcus

Johan D. Carlin[1], James B. Rowe[1,2,3], Nikolaus Kriegeskorte[1], Russell Thompson[1] and Andrew J. Calder[1]

[1]Medical Research Council Cognition and Brain Sciences Unit, Cambridge, CB2 7EF, UK, [2]Department of Clinical Neurosciences, University of Cambridge, Addenbrooke's Hospital, Cambridge, CB2 2QQ, UK and [3]Behavioural and Clinical Neuroscience Institute, University of Cambridge, Cambridge, CB2 3EB, UK

Address correspondence to email: johan.carlin@mrc-cbu.cam.ac.uk.

Humans and other primates are adept at using the direction of another's gaze or head turn to infer where that individual is attending. Research in macaque neurophysiology suggests that anterior superior temporal sulcus (STS) contains a direction-sensitive code for such social attention cues. By contrast, most human functional Magnetic resonance imaging (fMRI) studies report that posterior STS is responsive to social attention cues. It is unclear whether this functional discrepancy is caused by a species difference or by experimental design differences. Furthermore, social attention cues are dynamic in naturalistic social interaction, but most studies to date have been restricted to static displays. In order to address these issues, we used multivariate pattern analysis of fMRI data to test whether response patterns in human right STS distinguish between leftward and rightward dynamic head turns. Such head turn discrimination was observed in right anterior STS/superior temporal gyrus (STG). Response patterns in this region were also significantly more discriminable for head turn direction than for rotation direction in physically matched ellipsoid control stimuli. Our findings suggest a role for right anterior STS/STG in coding the direction of motion in dynamic social attention cues.

Keywords: face perception, fMRI, gaze, head, MVPA

Introduction

Humans and other primates share a remarkable ability to accurately perceive where other individuals are attending and use this information to change their own attentional state (Deaner and Platt 2003). Many higher order social cognitive processes depend on such gaze following behaviors (Frith and Frith 2008; Klein et al. 2009). Although changes to gaze direction and head view are inherently dynamic, to date the majority of human neuroimaging research has used static facial stimuli to study the neural representation of such social cues (Nummenmaa and Calder 2009). In view of macaque neurophysiology evidence that neurons responsive to dynamic head turns do not respond to static views of the same head (Hasselmo et al. 1989), it is vital to explore the neural coding of dynamic social stimuli. Here, we demonstrate that a region in human superior temporal sulcus (STS)/superior temporal gyrus (STG) contains a distributed representation of perceived head turn direction, thus supplying a necessary perceptual component to support a range of social behaviors.

Neurons in macaque anterior STS play a well-established role in representing the perceived direction of others' social attention cues, as conveyed by head orientation, gaze direction, and body posture (Perrett et al. 1982, 1992; Perrett, Smith, Potter, et al. 1985; Wachsmuth et al. 1994). However, these constitute only a minority of visually responsive STS neurons and are either spatially distributed (Hasselmo et al. 1989) or are organized into fine-grained patches well beyond the resolution of conventional functional MRI (fMRI; Perrett et al. 1984). This distributed representation poses a significant signal-to-noise challenge for attempts to study similar effects with human fMRI, where each voxel likely samples millions of neurons in ways that are only indirectly related to the neuronal spike trains commonly measured in macaque neurophysiology (Logothetis 2008; Kriegeskorte et al. 2009).

Unlike the typical anterior STS region identified by research in the macaque, most human fMRI studies report that social attention cues activate posterior STS and regions of adjacent STG and middle temporal gyrus (MTG; Hein and Knight 2008; Nummenmaa and Calder 2009). Similar posterior temporal regions are also more responsive to faces than to control stimuli (Andrews and Ewbank 2004; Fox et al. 2009). Most of these studies find that posterior STS is more responsive to averted than to direct gaze (Nummenmaa and Calder 2009), but the opposite pattern has also been observed (e.g., Pageler et al. 2003; Pelphrey et al. 2004). Furthermore, posterior STS responds more when an actor gazes away from a target than when the gaze direction is congruent with the target location (Pelphrey et al. 2003), suggesting that posterior STS is influenced by contextual effects, rather than by the direction of the social attention cue as such. Even in the absence of overt contextual manipulations, comparisons between direct and averted gaze may indirectly manipulate the engagement of approach/avoidance mechanisms and other higher order social cognitive functions associated with direct and averted gaze, such as theory of mind responses to eye contact (Emery 2000; Senju and Johnson 2009; Shepherd 2010). Thus, the litmus test for direction sensitivity is whether brain responses to different averted social attention cues can be distinguished in the absence of other contextual manipulations.

When such tests for direction sensitivity between different averted cues were carried out, one study found direction-sensitive fMRI adaptation to static images of gaze cues in right anterior, rather than posterior, STS (Calder et al. 2007). Another study that applied multivariate pattern analysis (MVPA) to a posterior STS region of interest (ROI) observed no distinction between different averted views of static heads (Natu et al. 2010) but did find that this ROI distinguished direct from averted head views across different head identities, suggesting an identity-invariant representation. These head view effects are consistent with the pattern of univariate sensitivity for direct against averted gaze observed in previous univariate research (Nummenmaa and Calder 2009). Considered collectively, this literature suggests a broad role for posterior STS in representing social attention cues, but unlike the evidence from macaque anterior STS, there is little indication that

posterior STS represents such cues in a direction-sensitive manner.

Outside the laboratory, cues to another's focus of attention are intrinsically dynamic in nature, but this issue has received limited attention in controlled experiments. There is initial evidence that a small subset of neurons in macaque anterior STS are tuned to dynamic changes in head turn direction (Perrett, Smith, Mistlin, et al. 1985; Hasselmo et al. 1989), but it remains unclear how the human brain codes such stimuli. In humans, posterior STS responds more to dynamic head turns than to both scrambled controls and static head views. However, neither anterior nor posterior STS has been found to show direction-sensitive coding of head turn direction, as measured by standard univariate fMRI (Lee et al. 2010). This absence of direction sensitivity is unsurprising, since neurons with such responses are unlikely to be clustered at a sufficiently large spatial scale to be detectable with univariate fMRI (Perrett et al. 1984; Hasselmo et al. 1989).

MVPA has recently been applied to detect representations thought to be coded in fine-grained patterns beyond the resolution of standard fMRI (Kamitani and Tong 2005; Haynes and Rees 2006; Shmuel et al. 2010). In the current study, we apply this method to determine whether distributed response patterns in the human STS region contain distinct direction-sensitive codes for observed head turns. If a classifier can use response patterns from the STS region to distinguish between leftward and rightward head turns, this would suggest that the underlying response patterns code head turn direction. However, leftward and rightward motion can also produce classification effects in regions without selectivity for social attention cues (Kamitani and Tong 2006). In order to avoid such confounding contributions of low-level motion, we included a set of rotating ellipsoid control videos. Previous work investigating head turn responses in macaque neurophysiology (Perrett, Smith, Mistlin, et al. 1985; Hasselmo et al. 1989) or direction-specific responses to static gaze (Calder et al. 2007) did not include such nonsocial controls, so an important aim of the current study was to establish that any direction-sensitive effects are specific to the social stimuli. Furthermore, we aimed to localize pattern effects to specific regions through the use of a searchlight algorithm that operated within the anatomically defined STS region. The STS region in this study included STG and MTG, in line with previous findings that social perception and gaze stimuli produce peaks that sometimes fall outside the STS proper (Allison et al. 2000; Nummenmaa and Calder 2009).

Materials and Methods

Participants

Twenty-one right-handed healthy volunteers with normal or corrected to normal vision participated in the study (12 males, mean age 29 years, age range 22–38). Volunteers provided informed consent as part of a protocol approved by the Cambridge Psychology Research Ethics Committee. Four volunteers were removed from further analysis: Two due to poor performance at the behavioral task whilst in the scanner (accuracy of less than 50%) and 2 due to fatigue and excessive head movements.

Experimental Design

Volunteers viewed 1000-ms video clips of 45° leftward and rightward head turns and comparable ellipsoid rotations (Fig. 1; e.g., videos, see Supplementary Material). Volunteers were instructed to monitor the stimulus set for infrequent deviant response trials (1 of the 8 experimental videos, rotated 4° from the upright position) and responded to detected deviants with a button press. The deviant response trials were drawn from all experimental conditions, and the degree of rotation was chosen after behavioral pilot tests to produce an attentionally demanding task without ceiling effects.

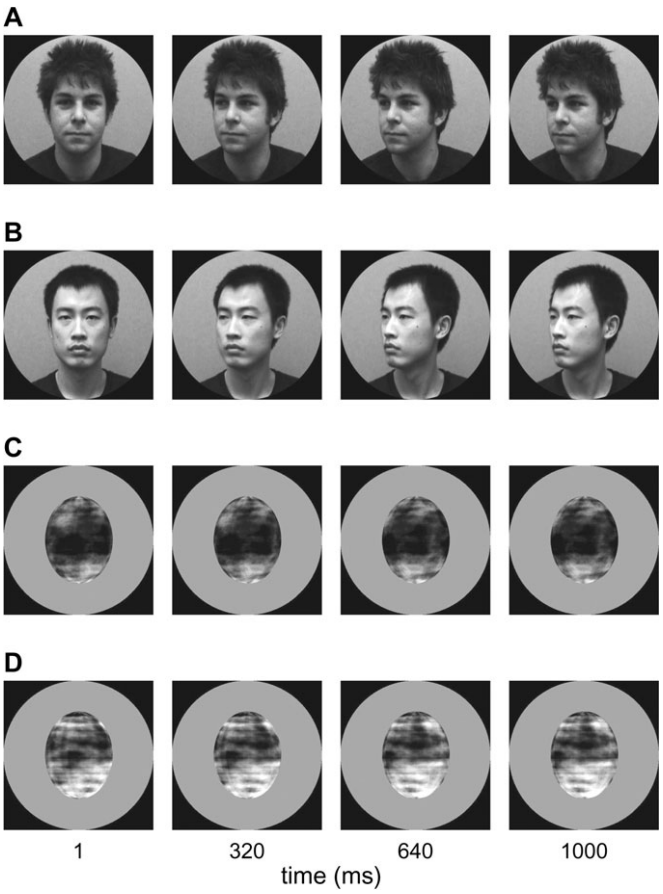


Figure 1. Example video frames for turning heads (*A–B*) and rotating ellipsoids (*C–D*). The stimuli were full color but are presented in grayscale for printing purposes (for full color stimuli, see Supplementary Videos). The videos were presented at 24 frames per second. All video frames are from leftward motion conditions. Rightward conditions were created through mirror reversal of the same video clips. The 2 ellipsoid identities (*C–D*) were created by Fourier-scrambling face textures from the first frame of the 2 head videos (*A–B*).

Two actors with matched head motion patterns were selected for the head turn videos. The ellipsoid control stimuli were rendered and animated in Matlab (Mathworks) and were texture mapped with the Fourier-scrambled face textures from the 2 head identities. The 2 motion directions were created by mirror reversing video clips with a single direction, thus ensuring that the stimulus set was physically matched across motion directions. This produced a total of 8 stimuli (2 heads, 2 ellipsoids, each rotating leftward or rightward), which were treated as individual conditions.

The stimuli were back-projected onto a screen in the scanner, which volunteers viewed via a tilted mirror. The stimuli were presented on a black background within a circular aperture (7° visual angle in diameter). The experiment was controlled using Matlab and the Psychophysics toolbox (Brainard 1997).

The experiment was divided into sets of 240 trials, each of which was independently randomized. Parameter estimates from each set formed an independent set of training examples for classification. The trials were presented within a rapid event-related design. Four volunteers completed a 6-set version of the experiment (approximately 40 min effective time) and 13 completed a 12-set version (80 min). Each set contained 240 trials: 80 null trials, where a fixation cross remained on

the screen throughout the trial (1500 ms) and 160 experimental trials (80 heads, 80 ellipsoids), where each trial consisted of a video clip (1000 ms) followed by fixation (500 ms). Each condition was repeated 18 times in a set. Sixteen deviant response trials were randomly sampled from the experimental conditions and responses to these trials were modeled with a separate nuisance regressor of no interest. The trials within the set were presented in a pseudorandomized order, where repeats of the same trial were slightly clustered in order to increase design efficiency (Henson 2003). Every second set was followed by a 15-s rest period, which was cued by a text prompt on the screen. The scan acquisition continued during the rest periods, and volunteers were instructed to remain still.

Imaging Acquisition

Scanning was carried out at the MRC Cognition and Brain Sciences Unit, Cambridge, United Kingdom, using a 3-T TIM Trio Magnetic Resonance Imaging scanner (Siemens), with a head coil gradient set. Functional data were collected using high-resolution echo planar T_2^*-weighted imaging (EPI, 40 oblique axial slices, time repetition [TR] 2490 ms, time echo [TE] 30 ms, in-plane resolution 2 × 2 mm, slice thickness 2 mm plus a 25% slice gap, 192 × 192 mm field of view). The acquisition window was tilted up approximately 30° from the horizontal plane to provide complete coverage of the occipital and temporal lobes. Preliminary pilot tests suggested that the use of this high-resolution EPI sequence resulted in reduced signal dropout in the anterior STS region, relative to a standard resolution sequence (3 × 3 × 3.75 mm voxels). All volumes were collected in a single continuous run for each volunteer. The initial 6 volumes from each run were discarded to allow for T_1 equilibration effects. T_1-weighted structural images were also acquired (magnetization prepared rapid gradient echo, 1 mm isotropic voxels).

Imaging Analysis

Imaging data were processed using statistical parametric mapping 5 (SPM5; www.fil.ion.ucl.ac.uk/spm). All functional volumes were realigned to the first nondiscarded volume, slice time corrected, and coregistered to the T_1 structural volume. The processing pathways for univariate analysis and MVPA diverged after these common steps (Fig. 2).

Univariate analysis was carried out using standard processing steps in SPM5. Structural volumes were segmented into gray and white matter partitions and spatially normalized to the Montreal Neurological Institute (MNI) template using combined segmentation and normalization routines. Functional volumes were normalized according to the parameters of this transformation, smoothed (10-mm full width at half mean Gaussian kernel, FWHM), and high pass filtered to remove low frequency drift (128-s cutoff period).

Subject-specific generalized linear models were used to analyze the data. The models included one regressor per condition and nuisance regressors for deviant response trials, volunteer responses to nondeviant trials, and for nulling scans that contained excessive noise or movement (Lemieux et al. 2007; Rowe et al. 2008; greater than 10 units intensity difference from the mean scaled image variance or more than 0.3 mm translational or 0.035 radians rotational movement relative to the previous volume). The volunteer-specific models included 0–135 such scan nulling regressors (mean 35). The experimental predictors were convolved with a canonical hemodynamic response function, and contrast images were generated based on the fitted responses. These contrast images were then entered into second-level permutation-based random effects models using statistical nonparametric mapping (SnPM; Nichols and Holmes 2002; 10 000 permutations, 10-mm FWHM variance smoothing).

Multivariate pattern analyses were carried out using functional volumes that had been realigned and slice timing corrected but had not been spatially normalized to the MNI template (Fig. 2). Each volunteer's data were modeled using a generalized linear model with similar regressors as in the univariate analysis, with the exception that each set of trials was modeled using a separate set of regressors. Individual parameter volumes from the first half of the data set was then averaged pairwise with the corresponding volume from the second half of the data set, thus reducing session effects at the expense of halving the number of training examples. This produced 3 or 6 final sets of examples to be used for classification, depending on the number of available sets before averaging. The example volumes were z-scored so that each voxel within a set had a mean of 0 and a standard deviation of 1 across examples in that set. Finally, each example was gray matter masked using the tissue probability maps generated by the segmentation processing stage.

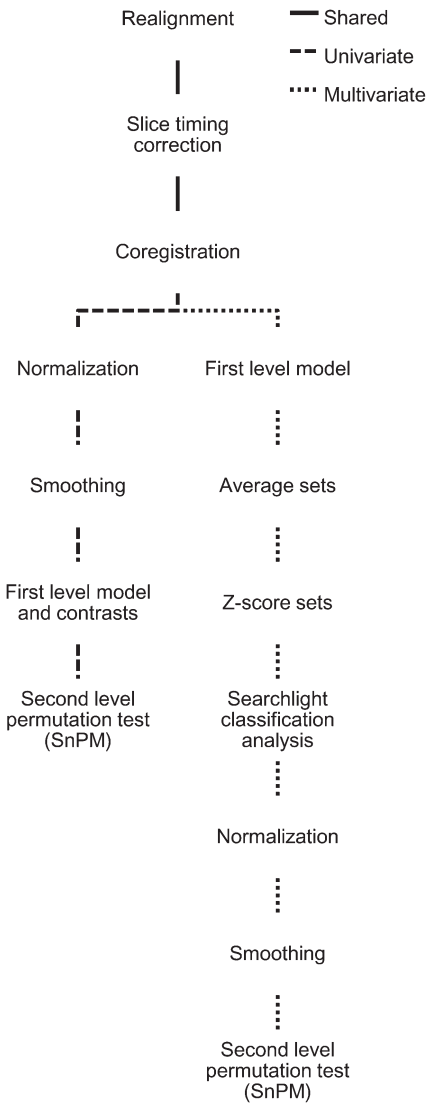


Figure 2. Processing pathways for fMRI analysis. All processing nodes take the result of the previous node in the hierarchy as input. With the exception of the searchlight classification analysis, all processing steps were implemented using standard SPM5 functionality.

The resulting example volumes were used as input to a linear support vector machine classifier (as implemented in PyMVPA; Hanke et al. 2009). All MVPA used a searchlight algorithm (Kriegeskorte et al. 2006), in which classification is carried out within a spherical region (5 mm radius) that is moved through the volume. Leave-one-out cross-validated classification accuracy estimates (percent correct) were mapped back to the center of each searchlight, thus producing a classification accuracy map.

The classification accuracy maps for each volunteer were normalized to MNI space, smoothed (10-mm FWHM), and entered into second-level nonparametric random effects models in SnPM. We used nonparametric tests because the discontinuous nature of the gray matter–masked data means that conventional familywise error (FWE) correction for multiple comparisons using random field theory in SPM5

would be inappropriate. We also wished to avoid making distributional assumptions about the first-level classification accuracy maps.

In line with the hypothesized site of our effects, we restricted our primary analysis to the right STS region, which was defined anatomically by the first author based on the mean T_1 volume for the sample. In line with previous evidence that social perception and eye gaze effects in the STS region extend into STG and MTG (Allison et al. 2000; Nummenmaa and Calder 2009), the mask included these gyri, whilst leaving out voxels in inferior temporal sulcus (inferior) or lateral fissure (superior) (Supplementary Fig. 1). We report P values corrected for multiple comparisons (FWE, $P < 0.05$) within this ROI (5162 voxels, y −58 to 22 mm MNI). We also carried out an exploratory analysis in a mirror-reversed version of the STS mask to test for effects in left STS. The use of a mirror-reversed mask sacrifices some anatomical precision in left STS but preserves the same voxel count and spatial structure in both masks. Visual inspection of the relation between the left STS mask and the mean T_1 volume suggested that the mask followed the anatomy of the sulcus in a comparable manner to the right STS mask. Finally, effects that survived correction for the full volume are also reported (FWE, $P < 0.05$). All analyses were restricted to a group gray matter mask, which was formed by the union of each volunteer's normalized individual gray matter mask. This mask ensured that we only considered effects in regions actually covered by the searchlight analysis.

Results

Behavioral Task

Volunteers were asked to detect the occasional 4° rotation of the video stimuli and were able to detect such deviant response trials adequately (mean accuracy 71%, standard error 4%). A repeated measures analysis of variance (ANOVA) of accuracy scores with the factors of stimulus type (head, ellipsoid) and motion direction (leftward, rightward) yielded no main effects and no interaction ($F_{1,16} < 2.4$, $P > 0.14$ for all effects), suggesting that volunteers did not assign attention differently to the heads and ellipsoids or to the 2 motion directions.

Multivariate Pattern Analysis

Superior Temporal Sulcus

Our primary hypothesis was that the right anterior STS region distinguishes between leftward and rightward perceived head turns. In line with this prediction, a group analysis of the MVPA searchlight results for the right STS region showed that classification of head turn direction was significantly more accurate than expected by chance in a right anterior STS/STG site ($P = 0.005$ FWE, 50 4 −14 mm MNI, Fig. 3*A*; for individual subject results, see Supplementary Fig. 2). By comparison, left–right classification of rotation direction in the ellipsoid control stimuli exceeded chance in middle STS ($P = 0.037$ FWE, 50 −14 −10 mm MNI, Fig. 3*B*).

The peaks of these head turn and ellipsoid rotation effects were approximately 18 mm apart and the activated regions did not overlap, which raises the question of how distinct the 2 effects are. We addressed this by computing the difference between the classification maps for head turn and ellipsoid rotation in each volunteer. These difference maps were entered into a group analysis, which showed that left–right classification was more accurate for head turns than for ellipsoid rotations in right anterior STS/STG ($P = 0.027$ FWE, 52 12 −12 mm MNI, Fig. 3*C*). This effect overlapped with the head turn classification effect (8 mm distance between peaks, 40% overlap), suggesting a common origin. No STS region showed significantly more accurate direction classification for ellipsoid rotations than for head turns.

We tested whether the left–right head turn codes were invariant to head identity by training the classifier on the left–right turns of one head and applying the learned weights to left–right turns of the other head. Left–right classification did not generalize across head identity at any site in right STS. Similarly, there was no significant left–right generalization

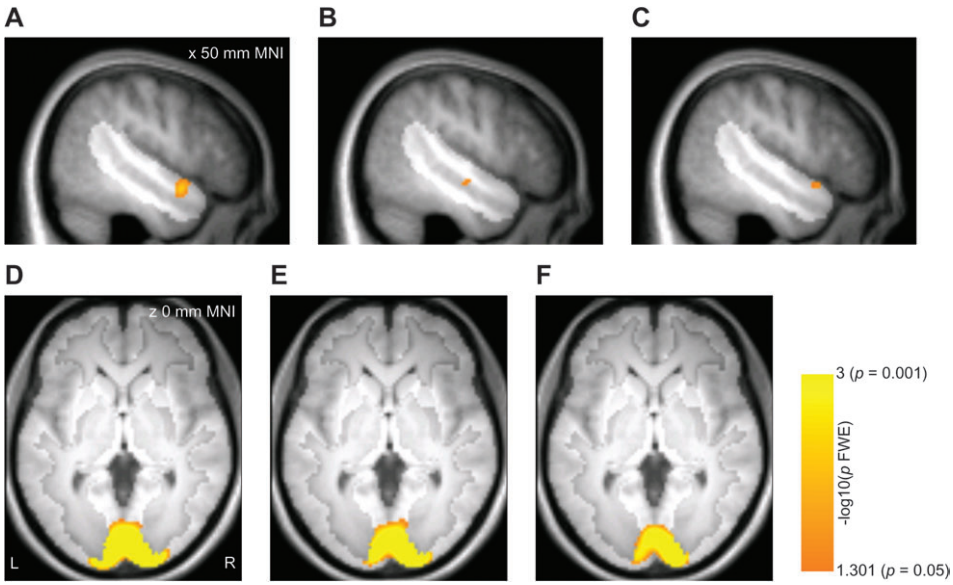


Figure 3. Group results for MVPA, displayed on the mean T_1 volume for the sample. Effects are displayed corrected for multiple comparisons within the right STS region (panels *A–C*; hypothesis-driven analysis, $P < 0.05$ FWE) or the full gray matter volume (panels *D–F*; exploratory analysis, $P < 0.05$ FWE). The highlighted portion of each panel shows the extent of the mask. (*A*) Classification of left–right head turns in the right STS/STG region. (*B*) Classification of left–right ellipsoid rotations in the right STS region. (*C*) Right STS regions where left–right classification of head turns was more accurate than classification of ellipsoid rotations. (*D*) Classification of left–right head turns in the full gray matter volume. (*E*): Gray matter regions where left–right classification of head turns was more accurate than classification of ellipsoid rotations. (*F*) Gray matter regions where the weights acquired by training the classifier on left–right head turns for one head identity generalized to left–right head turns in the other head identity.

across ellipsoid identities and no left–right generalization across stimulus type (head and ellipsoid).

We also carried out an exploratory analysis of effects in the left anatomically defined STS region. No left STS regions showed above-chance classification of observed head turn direction. However, a region in left anterior STS distinguished ellipsoid rotation direction with above-chance accuracy (P = 0.01 FWE, –56 –8 –16 mm MNI, Supplementary Fig. 3*A*). Direction classification accuracy was significantly higher for ellipsoid rotation than for head turns in a similar region (P = 0.041 FWE, –58 –4 –16 mm MNI, Supplementary Fig. 3*B*). No other classification effects were significant in this ROI.

Whole-Brain Analysis

Beyond our hypothesis-driven search within the anatomically defined right STS region, we also carried out an exploratory analysis within the full gray matter–masked volume to identify other effects of interest. Classification of left–right head turns exceeded chance in a region including calcarine sulcus and occipital pole (P < 0.001 FWE, 16 –96 0 mm MNI, Fig. 3*D*). This region is likely to include visual areas V1, V2, and V3, but in the absence of a retinotopic localizer, we use the general term early visual cortex to describe this region. Left–right ellipsoid classification did not produce significant effects in any region. Left–right classification was significantly more accurate for head turns than for ellipsoid rotations in a similar early visual region (P < 0.001 FWE, 14 –96 2 mm MNI, Fig. 3*E*). A similar region in early visual cortex also allowed left–right classification to generalize across head identities (P < 0.001 FWE, 14 –96 2 mm MNI, Fig. 3*F*) but not across stimulus types. No regions outside of early visual cortex showed significant effects for any of these comparisons.

Univariate Analysis

We used a conventional univariate analysis in SPM5 to address whether the observed classification effects could be attributed to large-scale response level differences between the conditions. To make comparisons between MVPA and univariate results simpler, the univariate analysis also used nonparametric permutation-based random effects analysis of group effects (SnPM, for details, see Materials and Methods). We also explored whether direction classification of head turns colocalized with greater univariate responses to heads than to ellipsoids.

Superior Temporal Sulcus

No regions inside the anatomically defined right STS ROI responded selectively to one head turn direction over the other or to one ellipsoid rotation direction over the other, suggesting that the left–right classification effects in this region did not co-occur with large-scale univariate direction sensitivity.

Collapsing across motion direction, right posterior STS responded significantly more to heads than to ellipsoids (P < 0.002 FWE, 48 –44 16 mm MNI, Fig. 4*A*), while a region in middle STG bordering on the edge of the ROI responded more to ellipsoids than to heads (P = 0.004 FWE, 60 0 0 mm MNI, Fig. 4*B*). Thus, univariate selectivity for heads over ellipsoids occurred in posterior STS, 57 mm from the left–right head turn classification peak in anterior STS/STG. The peaks for univariate selectivity for ellipsoids over heads and for left–right ellipsoid rotation classification were separated by 20 mm.

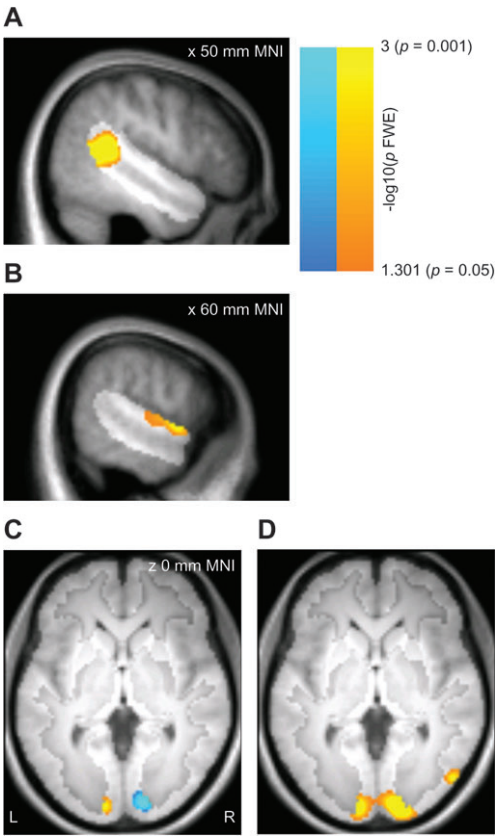


Figure 4. Group results for the univariate analysis, displayed on the mean T_1 volume for the sample. Effects are displayed corrected for multiple comparisons within the right STS region (panels *A–B*; hypothesis-driven analysis, P < 0.05 FWE) or the full gray matter volume (panels *C–D*; exploratory analysis, P < 0.05 FWE). The highlighted portion of each panel shows the extent of the mask. (*A*) Greater univariate responses to heads than to ellipsoids in the right STS region. (*B*) Greater univariate responses to ellipsoids than to heads in the right STS region. (*C*) Gray matter regions with greater univariate responses to left than to right head turns (warm colors) or with greater univariate responses to right than to left head turns (cool colors). The effects do not overlap at any site. (*D*) Gray matter regions with greater univariate responses to heads than to ellipsoids.

Neither of the univariate effects overlapped with the classification effects.

Within the left STS ROI, a posterior region responded more to heads than to ellipsoids (P = 0.004 FWE, –52 –58 14 mm MNI, Supplementary Fig. 3*C*) and left middle STS responded more to ellipsoids than to heads (P = 0.014, –66 –18 –14 mm MNI, Supplementary Fig. 3*D*), mirroring the results obtained in the right STS region. No left STS regions responded preferentially to head turn or ellipsoid rotation in one direction relative to another. No other comparisons reported above were significant in the left STS analysis.

Whole-Brain Analysis

A univariate analysis of the gray matter–masked full volume revealed significant univariate selectivity for left over right head turns that was restricted to left early visual cortex (P < 0.001 FWE, –12 –94 0 mm MNI), and conversely, selectivity for right over left head turns restricted to right early visual cortex (P < 0.001 FWE, 14 –92 4 mm MNI, Fig. 4*C*). These effects almost completely overlapped the left–right head turn classification effect in early visual cortex (100% overlap for left over right, 91% overlap for right over left), suggesting that the

classification effects co-occurred with large-scale univariate effects. Note that the laterality of these early visual effects is opposite to what would be expected for a stimulus that moves into the right and left visual hemifields, a point we return to below. No regions showed a preference for one ellipsoid rotation direction over the other in the whole-brain analysis.

A comparison of univariate responses to heads over ellipsoids and ellipsoids over heads revealed a network of activations (Supplementary Table 1). Of primary interest to the current study, bilateral early visual cortex responded more to heads than to ellipsoids ($P < 0.001$ FWE, 18 –96 –4 mm MNI, Fig. 4*D*), and this early visual effect overlapped the left–right head turn classification effect (91% overlap). Thus, the left–right head turn classification effects occurred in a region where we also observed univariate selectivity for head turn direction and preferential responses to heads over ellipsoids. Bilateral regions in posterior MTG also responded more to heads than to ellipsoids (right: $P = 0.001$ FWE, 52 –74 2 mm MNI. Left: $P = 0.001$ FWE, –50 –72 14 mm MNI). These coordinates are close to those previously reported for motion area MT (Dumoulin et al. 2000; conversion from Talairach to MNI coordinates with tools from Evans et al. 2007). Because we did not include a specific localizer scan to distinguish MT from MST or other motion areas, we refer to this region as MT+. The MT+ regions showed no direction-sensitive responses in the univariate or classification analyses, even at reduced thresholds ($P < 0.01$, uncorrected).

Follow-up Experiments

The pattern of univariate effects in early visual cortex suggested to us the presence of eye movements in the experiment. If volunteers tracked the heads as they turned, this would have placed the stimulus primarily in the hemifield ipsilateral to the direction of motion, which could explain the ipsilateral univariate activations in early visual cortex. Eye tracking was not available when the main experiment was undertaken, so we carried out follow-up eye tracking and fMRI experiments with 3 principal aims: First, to test whether the head turns used in the main experiment elicit eye movements; second, to assess whether the eye movement effects could be removed with a revised experimental paradigm; and finally, to test whether the fMRI effects reported in the main text remained in the absence of statistically significant eye movement effects.

Follow-up Materials and Methods

Five volunteers from the final sample used in the main experiment returned to participate in additional experiments. Eye movements were monitored using a video-based infrared eye tracker (500 Hz acquisition outside the scanner, 50 Hz acquisition inside the scanner; Sensomotoric Instruments). We analyzed the change in horizontal fixation position at the end relative to the start of each trial using custom code developed in Matlab.

Imaging data were acquired and analyzed using identical parameters as in the main experiment, with the exception that no averaging of the first and second half of the experiment was carried out, since this would have yielded an unacceptably small number of observations for first-level statistics. Furthermore, each set was scanned in a separate run to allow recalibration of the eye tracker between sets. As in the main fMRI experiment, we used a searchlight analysis. We based single-volunteer inference on binomial tests at each voxel in the ROI.

Follow-up Eye Tracking with the Original Design

Five volunteers carried out an abbreviated version of the main experiment outside the scanner (3 sets, 540 trials), while their eye position was monitored. First-level ANOVAs revealed that each volunteer showed a significant stimulus type (head, ellipsoid) by motion direction (leftward, rightward) interaction (Supplementary Table 2). This interaction reflected consistent fixation shifts in the direction of the head turns, with nonsignificant or weaker fixation shifts in the direction of the ellipsoid rotations.

Follow-up Eye Tracking with Revised Design

We carried out a second eye tracking experiment with a revised paradigm that included a fixation cross during the presentation of the video clips. Volunteers were also strongly instructed to maintain fixation at all times. We included only the head turn conditions in order to obtain a maximal number of trials for the head left–right comparison whilst minimizing volunteer fatigue. In this second experiment, the head turn left–right eye movement effect was reduced to nonsignificance in 4 of 5 volunteers (Fig. 5*A–C*).

Follow-up fMRI Experiment with Revised Design

We tested whether our main classification findings in STS/STG and early visual cortex survived in the absence of eye movements by carrying out a second fMRI experiment with the revised experimental paradigm. We recruited the 4 volunteers who showed no significant eye movement effects in the eye tracking test outside the scanner. Volunteers completed a full 6-set version of the revised experiment (1080 trials, for details, see Material and Methods), while their eye position was monitored. One of the 4 scanned volunteers showed a significant fixation shift in response to the head turns whilst being scanned ($t = 8.72$, $P = 0.003$). This volunteer was removed from further analysis.

Although the 3 remaining volunteers showed no significant eye movement effects (as observed in separate tests before and during scanning), left–right classification of head turns in the right anterior STS region was greater than chance in 2 volunteers ($P < 0.05$, Bonferroni FWE corrected for the right STS mask) and at reduced thresholds in the third ($P < 0.001$, uncorrected, Fig. 5*D–F*). The final volunteer also showed an effect in posterior STS ($P < 0.05$, FWE).

All 3 volunteers showed significant left–right head turn classification effects in early visual cortex ($P < 0.05$ Bonferroni FWE corrected for a 20 mm radius sphere centered on the peak head turn classification effect in the main experiment, Fig. 5*G–I*). However, unlike the main experiment, where this effect was joined by univariate response preferences for head turns in a direction ipsilateral to the visual hemifield (Supplementary Fig. 4*A*), we now observed preferentially contralateral responses to head turn direction ($P < 0.001$, uncorrected, Supplementary Fig. 4*B*). Thus, although the classification effects in early visual cortex were accompanied by univariate effects also in the revised experiment, laterality of these univariate effects was reversed.

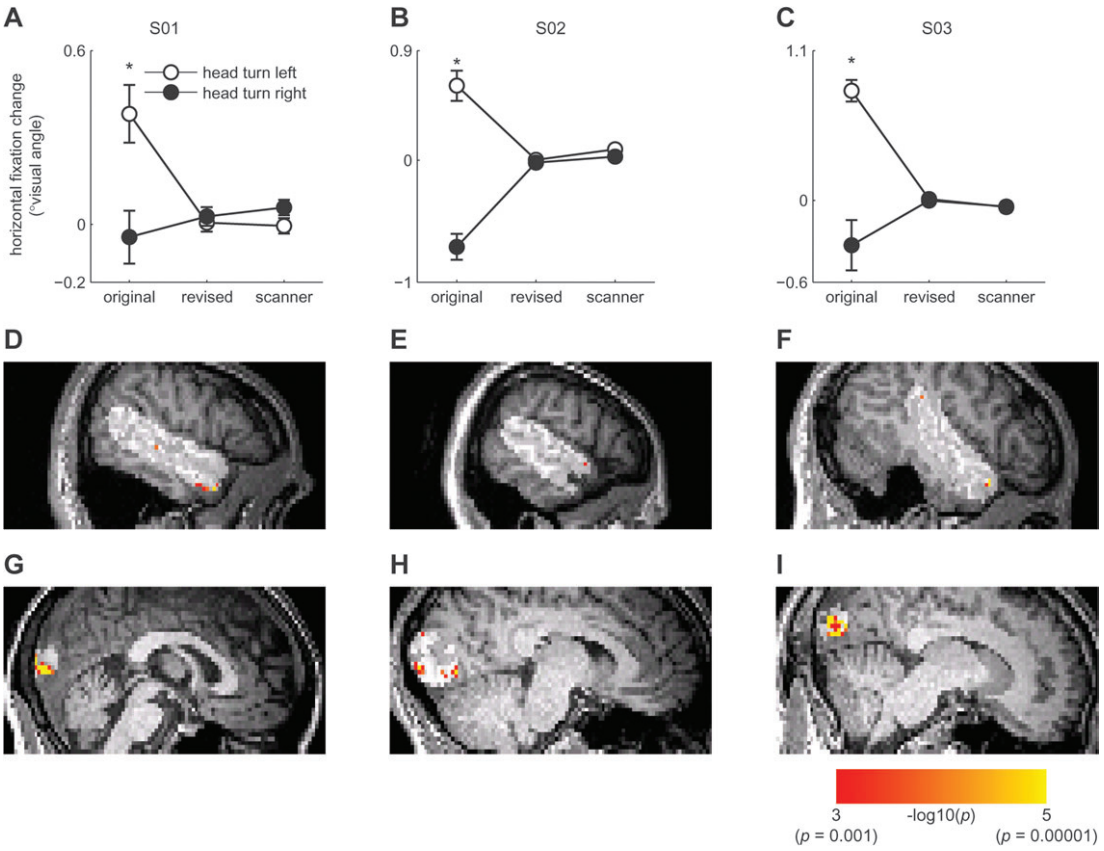


Figure 5. Follow-up eye tracking and fMRI experiments. (*A–C*) Mean horizontal fixation change plotted separately for the 3 volunteers selected for the final analysis in the revised fMRI experiment. Positive values reflect a leftward shift in fixation over the trial, while negative values reflect a rightward shift. The horizontal axis gives fixation performance in the original task, the revised task, and the revised task as measured during the fMRI experiment. The error bars give ±1 standard error of the mean. Comparisons with significant differences between the head turn directions are highlighted by asterisks (*t*-tests, $P < 0.05$). It can be seen that the revised design abolished the eye movement effect in these volunteers. (*D–F*) Left–right head turn classification results for the 3 volunteers in the final sample of the fMRI experiment. The volunteers are shown in the same order as in *A–C*. Results are overlaid on each volunteer's T_1 volume and are masked to only include effects within the highlighted right STS region ($P < 0.001$, uncorrected). It can be seen that even in the absence of eye movement effects, anterior STS/STG codes head turn direction. (*G–I*) Results as in *D–F* but masked to show effects within a 20 mm radius of the peak early visual head turn classification effect from the main study. It can be seen that the effects in early visual cortex also remain when eye movements are controlled.

Discussion

Appropriate social behavior is dependent on accurately inferring where others are attending. In the visual domain, this inferential process is likely to involve direction-sensitive coding of social attention cues, such as head turns. In experiments, these stimuli are often abstracted to static views, which fails to capture their dynamic character in natural social interaction. Here, we demonstrate that response patterns in human right anterior STS/STG distinguish between leftward and rightward dynamic head turns. Furthermore, left–right head turns were significantly more discriminable in this region than left–right ellipsoid control stimuli. A similar analysis of the left STS region revealed no left–right classification of head turn direction at any site in the ROI.

The peak coordinates for left–right classification of head turn direction in the current study are in close proximity to a previous demonstration of direction-sensitive fMRI adaptation to static gaze (Calder et al. 2007; 16 mm distance between peaks). Considered collectively, these results suggest a general role for right anterior STS/STG in supplying higher order social cognitive processes with important information about the direction of another's attentional shifts, whether these are conveyed by static gaze in a front-facing head or dynamic head turns. Consistent with this social role, we also demonstrate that direction sensitivity does not extend to nonsocial control stimulus motion in this region. An important question is whether such direction-sensitive responses to dynamic and static social cues are driven by a single representation of the direction of another's social attention (Perrett et al. 1992) or whether dynamic information is coded separately, as indicated by the finding that STS neurons tuned to head turn motion do not respond to static head view displays (Perrett, Smith, Mistlin, et al. 1985; Hasselmo et al. 1989).

Neurons in macaque anterior STS are tuned to the direction of social attention cues (Hasselmo et al. 1989; Perrett et al. 1992). However, most human fMRI studies have reported gaze or head turn effects in posterior rather than anterior STS regions (Nummenmaa and Calder 2009). Our classification effects appear more consistent with the typical recording site in macaque anterior STS than with previous univariate fMRI effects in human posterior STS. Compared with standard univariate analysis, MVPA and fMRI adaptation techniques confer greater sensitivity (Haynes and Rees 2006). This increased sensitivity makes more rigorous comparisons possible, for instance between left and right averted social attention cues. Accordingly, we also observed greater consistency between human fMRI and single unit evidence from the

macaque (see also Kamitani and Tong 2005, 2006; Calder et al. 2007). Known human–macaque discrepancies in the function of posterior STS and surrounding areas suggest that a simple correspondence between human and macaque may not apply to all high-level visual areas (Orban et al. 2004), but such a simple correspondence nevertheless offers a useful working model for the representation of social attention cues.

The pattern of results we observed in posterior and anterior portions of the right STS region also highlights how large-scale univariate response level differences can dissociate from multivariate classification performance (Haxby et al. 2001; Hanson et al. 2004; Hanson and Halchenko 2008). Similar to previous studies (Andrews and Ewbank 2004), we found that right posterior STS responded more to heads than to ellipsoids, while no such preferential responding was observed in anterior STS/STG. The left–right head turn classification effects showed the opposite pattern, with significant effects in anterior but not posterior regions. There are clear parallels between this pattern of effects and a recent report where face identity classification was possible in an anterior inferotemporal region, which did not respond preferentially to faces over places, while no such face identity effects appeared in the more posterior fusiform face area, even though this region responded more to faces than to places (Kriegeskorte et al. 2007). Face identity and head turn direction are both important dimensions for face processing, yet multivariate sensitivity for manipulations along these dimensions does not appear to colocalize with univariate selectivity for faces over other object categories. Although more systematic studies of these within- and between-category dissociations are needed before their theoretical implications for face perception can be fully considered, the current results indicate that studies where data analysis is restricted to functional ROIs defined by face selectivity are at risk of missing potentially important effects (Haxby et al. 2001; Friston et al. 2006).

Neurons with social attention responses in macaque STS are often invariant to the identity of the individual conveying the cue (Perrett et al. 1992). In this study, we observed no generalization between response patterns evoked by left–right head turns across the 2 identities. Although there is some initial evidence to suggest that STS neurons can code both head view and head identity (Perrett et al. 1984), it is in our view unlikely that the representation across STS is identity-specific. For instance, it has previously been shown that direct and averted static head views can be distinguished across identity in posterior STS (Natu et al. 2010). Given that separate training of left–right classification for each identity involves half as much data as compared with when this dimension is collapsed, it is more likely that our experiment was not sufficiently sensitive to detect any such identity-invariant head turn representations.

Our results suggest that the anterior STS region distinguishes the direction of perceived head turns. The follow-up eye-tracking experiment suggested that volunteers' eye movements tended to follow the direction of head turns, thus presenting a potential confound to the interpretation of our results. To rule out an eye movement account of our reported classification effects, we demonstrated in a revised fMRI experiment that a subset of volunteers from the main experiment showed significant left–right head turn classification in the right STS region, even though these volunteers showed no significant eye movement effects during pretests or whilst in the scanner. Thus, even though our main analysis is potentially limited by an eye movement confound, the head turn direction codes in the right anterior STS region remain when this confound is removed. The absence of prior reports of eye movement responses in the anterior STS region is also consistent with this interpretation (Grosbras et al. 2005; Bakola et al. 2007). By contrast, even minute eye movements elicit responses in early visual cortex (Dimigen et al. 2009), and an eye movement account would seem to account well for the pattern of ipsilateral univariate selectivity we observed in the main experiment, with leftward and rightward head turns producing responses in left and right early visual cortex, respectively. Notably, this ipsilateral pattern of effects reverted to the expected contralateral response preference in the univariate analysis of the follow-up experiment, even though left–right head turn classification in early visual cortex was significant in both the original and the follow-up experiments. These results suggest that the classification effects in the 2 data sets were driven by distinct large-scale univariate effects: a primarily eye movement-related response in the main experiment and a visually-evoked response in the follow-up experiment.

The pervasive tendency for volunteers to follow social attention cues points to an intriguingly close link between action and perception in this system, which is worthy of further enquiry. Previous investigators found that static gaze cues also evoke small eye movements in the perceived gaze direction (Mansfield et al. 2003). Indeed, 2 of the 5 volunteers who were tested with eye tracking in the current study were unable to consistently suppress eye movements in response to the head turns, even in the presence of a fixation cross and strong instructions to maintain fixation. Although interesting in their own right, these eye movement effects also suggest that investigators who seek to isolate effects of perceived gaze direction would be well advised to monitor the volunteer's own gaze.

Previous studies have found that socially relevant motion engages MT (Puce et al. 1998; Watanabe et al. 2006). Consistent with this literature, we observed a univariate response preference for heads relative to ellipsoids in bilateral superior temporal regions likely corresponding to MT+. Despite this category preference for heads relative to ellipsoids, we obtained no evidence that response patterns in this region distinguish head turn direction. In previous studies that attempted to decode motion directions, direction sensitivity was weaker in MT than in earlier visual areas (Kamitani and Tong 2006; Seymour et al. 2009), which the authors attribute to MT's smaller anatomical size compared with earlier visual areas. Although neurophysiological data suggest considerable direction sensitivity in both MT and early visual cortex (Snowden et al. 1992), such response properties may interact with area size when measured with coarse-grained methods such as fMRI, thus producing apparently weaker or nonsignificant effects in smaller areas (Bartels et al. 2008). Note also that both the absence of a functional MT localizer and the use of weaker, more transient motion stimuli may have rendered our analysis less sensitive to direction-sensitive responses in MT+, compared with previous studies (Kamitani and Tong 2006; Seymour et al. 2009). Thus, we do not exclude the possibility that head turns produce direction-sensitive MT+ responses, although we were unable to find evidence for this.

In conclusion, we have presented evidence that response patterns in human right anterior STS/STG distinguish between leftward and rightward dynamic head turns. Such direction sensitivity was not detected for physically matched ellipsoid control stimuli. The anterior site of this effect is consistent with

evidence from macaque neurophysiology (Perrett, Smith, Mistlin, et al. 1985; Hasselmo et al. 1989) but does not colocalize with regions showing greater univariate responses to heads than to ellipsoids. In this respect, multivariate pattern approaches show great promise in linking evidence from single neurons in the macaque to large-scale response patterns in human fMRI.

Supplementary Material

Supplementary material can be found at: http://www.cercor.oxfordjournals.org/

Funding

United Kingdom Medical Research Council (grant MC_US_A060_0017 to A.J.C.; MC_US_A060_0016 to J.B.R); Wellcome Trust (WT077029 to J.B.R.).

Notes

We are grateful to Christian Schwarzbauer for help with developing the high-resolution EPI sequence used in this study. *Conflict of Interest*: None declared.

References

Allison T, Puce A, McCarthy G. 2000. Social perception from visual cues: role of the STS region. Trends Cogn Sci. 4:267–278.

Andrews T, Ewbank M. 2004. Distinct representations for facial identity and changeable aspects of faces in the human temporal lobe. Neuroimage. 23:905–913.

Bakola S, Gregoriou G, Moschovakis A, Raos V, Savaki H. 2007. Saccade-related information in the superior temporal motion complex: quantitative functional mapping in the monkey. J Neurosci. 27:2224–2229.

Bartels A, Logothetis NK, Moutoussis K. 2008. fMRI and its interpretations: an illustration on directional selectivity in area V5/MT. Trends Neurosci. 31:444–453.

Brainard D. 1997. The psychophysics toolbox. Spat Vis. 10:433–436.

Calder AJ, Beaver JD, Winston JS, Dolan RJ, Jenkins R, Eger E, Henson RNA. 2007. Separate coding of different gaze directions in the superior temporal sulcus and inferior parietal lobule. Curr Biol. 17:20–25.

Deaner R, Platt M. 2003. Reflexive social attention in monkeys and humans. Curr Biol. 13:1609–1613.

Dimigen O, Valsecchi M, Sommer W, Kliegl R. 2009. Human microsaccade-related visual brain responses. J Neurosci. 29:12321–12331.

Dumoulin SO, Bittar RG, Kabani NJ, Baker CL, Le Goualher G, Pike GB, Evans AC. 2000. A new anatomical landmark for reliable identification of human area V5/MT: a quantitative analysis of sulcal patterning. Cereb Cortex. 10:454–463.

Emery NJ. 2000. The eyes have it: the neuroethology, function and evolution of social gaze. Neurosci Biobehav Rev. 24:581–604.

Evans A, Zilles K, Lancaster J, Martinez M, Mazziotta J, Fox P, Tordesillas-Gutierrez D, Salinas F. 2007. Bias between MNI and Talairach coordinates analyzed using the ICBM-152 brain template. Hum Brain Mapp. 28:1194–1205.

Fox C, Iaria G, Barton J. 2009. Defining the face processing network: optimization of the functional localizer in fMRI. Hum Brain Mapp. 30:1637–1651.

Friston K, Rotshtein P, Geng J, Sterzer P, Henson R. 2006. A critique of functional localisers. Neuroimage. 30:1077–1087.

Frith C, Frith U. 2008. Implicit and explicit processes in social cognition. Neuron. 60:503–510.

Grosbras M-H, Laird A, Paus T. 2005. Cortical regions involved in eye movements, shifts of attention, and gaze perception. Hum Brain Mapp. 25:140–154.

Hanke M, Halchenko Y, Sederberg P, Olivetti E, Fründ I, Rieger J, Herrmann C, Haxby J, Hanson SJ, Pollmann S. 2009. PyMVPA: a unifying approach to the analysis of neuroscientific data. Front Neuroinformatics. 3:1–13.

Hanson SJ, Halchenko Y. 2008. Brain reading using full brain support vector machines for object recognition: there is no "face" identification area. Neural Comput. 20:486–503.

Hanson SJ, Matsuka T, Haxby JV. 2004. Combinatorial codes in ventral temporal lobe for object recognition: Haxby (2001) revisited: is there a "face" area? Neuroimage. 23:156–166.

Hasselmo M, Rolls E, Baylis G, Nalwa V. 1989. Object-centered encoding by face-selective neurons in the cortex in the superior temporal sulcus of the monkey. Exp Brain Res. 75:417–429.

Haxby J, Gobbini M, Furey M, Ishai A, Schouten J, Pietrini P. 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science. 293:2425–2430.

Haynes J, Rees G. 2006. Decoding mental states from brain activity in humans. Nat Rev Neurosci. 7:523–534.

Hein G, Knight R. 2008. Superior temporal sulcus—it's my area: or is it? J Cogn Neurosci. 20:2125–2136.

Henson RNA. 2003. Analysis of fMRI timeseries: linear time-invariant models, event-related fMRI and optimal experimental design. In: Frackowiak RSJ, Friston KJ, Frith CD, Dolan RJ, Price CJ, editors. Human brain function. New York: Academic Press. p. 793–822.

Kamitani Y, Tong F. 2005. Decoding the visual and subjective contents of the human brain. Nat Neurosci. 8:679–685.

Kamitani Y, Tong F. 2006. Decoding seen and attended motion directions from activity in the human visual cortex. Curr Biol. 16:1096–1102.

Klein J, Shepherd S, Platt M. 2009. Social attention and the brain. Curr Biol. 19:R958–R962.

Kriegeskorte N, Bandettini P, Cusack R. 2009. How does an fMRI voxel sample the neuronal activity pattern: compact-kernel or complex-spatiotemporal filter? Neuroimage. 49:1965–1976.

Kriegeskorte N, Formisano E, Sorger B, Goebel R. 2007. Individual faces elicit distinct response patterns in human anterior temporal cortex. Proc Natl Acad Sci U S A. 104:20600–20605.

Kriegeskorte N, Goebel R, Bandettini P. 2006. Information-based functional brain mapping. Proc Natl Acad Sci U S A. 103:3863–3868.

Lee LC, Andrews T, Johnson SJ, Woods W, Gouws A, Green GGR, Young AW. 2010. Neural responses to rigidly moving faces displaying shifts in social attention investigated with fMRI and MEG. Neuropsychologia. 48:477–490.

Lemieux L, Salekhaddadi A, Lund T, Laufs H, Carmichael D. 2007. Modelling large motion events in fMRI studies of patients with epilepsy. Magn Reson Imaging. 25:894–901.

Logothetis N. 2008. What we can do and what we cannot do with fMRI. Nature. 453:869–878.

Mansfield EP, Farroni T, Johnson MH. 2003. Does gaze perception facilitate overt orienting? Vis Cogn. 10:7–14.

Natu VS, Jiang F, Narvekar A, Keshvari S, Blanz V, O'Toole AJ. 2010. Dissociable neural patterns of facial identity across changes in viewpoint. J Cogn Neurosci. 22:1570–1582.

Nichols T, Holmes A. 2002. Nonparametric permutation tests for functional neuroimaging: a primer with examples. Hum Brain Mapp. 15:1–25.

Nummenmaa L, Calder A. 2009. Neural mechanisms of social attention. Trends Cogn Sci. 13:135–143.

Orban GA, Van Essen D, Vanduffel W. 2004. Comparative mapping of higher visual areas in monkeys and humans. Trends Cogn Sci. 8:315–324.

Pageler NM, Menon V, Merin NM, Eliez S, Brown WE, Reiss AL. 2003. Effect of head orientation on gaze processing in fusiform gyrus and superior temporal sulcus. Neuroimage. 20:318–329.

Pelphrey K, Singerman J, Allison T, McCarthy G. 2003. Brain activation evoked by perception of gaze shifts: the influence of context. Neuropsychologia. 41:156–170.

Pelphrey K, Viola R, Mccarthy G. 2004. When strangers pass: processing of mutual and averted social gaze in the superior temporal sulcus. Psychol Sci. 15:598–603.

Perrett DI, Hietanen J, Oram M, Benson P. 1992. Organization and functions of cells responsive to faces in the temporal cortex. Philos Trans R Soc Lond B Biol Sci. 335:23–30.

Perrett DI, Rolls E, Caan W. 1982. Visual neurones responsive to faces in the monkey temporal cortex. Exp Brain Res. 47:329–342.
Perrett DI, Smith P, Potter D, Mistlin A, Head A, Milner A, Jeeves M. 1985. Visual cells in the temporal cortex sensitive to face view and gaze direction. Proc R Soc Lond B Biol Sci. 223:293–317.
Perrett DI, Smith PA, Mistlin AJ, Chitty AJ, Head AS, Potter DD, Broennimann R, Milner AD, Jeeves MA. 1985. Visual analysis of body movements by neurones in the temporal cortex of the macaque monkey: a preliminary report. Behav Brain Res. 16:153–170.
Perrett DI, Smith PA, Potter DD, Mistlin AJ, Head AS, Milner AD, Jeeves MA. 1984. Neurones responsive to faces in the temporal cortex: studies of functional organization, sensitivity to identity and relation to perception. Hum Neurobiol. 3:197–208.
Puce A, Allison T, Bentin S, Gore JC, McCarthy G. 1998. Temporal cortex activation in humans viewing eye and mouth movements. J Neurosci. 18:2188–2199.
Rowe J, Eckstein D, Braver T, Owen A. 2008. How does reward expectation influence cognition in the human brain? J Cogn Neurosci. 20:1980–1992.
Senju A, Johnson M. 2009. The eye contact effect: mechanisms and development. Trends Cogn Sci. 13:127–134.
Seymour K, Clifford C, Logothetis N, Bartels A. 2009. The coding of color, motion, and their conjunction in the human visual cortex. Curr Biol. 19:177–183.
Shepherd SV. 2010. Following gaze: gaze-following behavior as a window into social cognition. Front Integr Neurosci. 4:1–5.
Shmuel A, Chaimow D, Raddatz G, Ugurbil K, Yacoub E. 2010. Mechanisms underlying decoding at 7 T: ocular dominance columns, broad structures, and macroscopic blood vessels in V1 convey information on the stimulated eye. Neuroimage. 49:1957–1964.
Snowden R, Treue S, Andersen R. 1992. The response of neurons in areas V1 and MT of the alert rhesus monkey to moving random dot patterns. Exp Brain Res. 88:389–400.
Wachsmuth E, Oram M, Perrett DI. 1994. Recognition of objects and their component parts: responses of single units in the temporal cortex of the macaque. Cereb Cortex. 4:509–522.
Watanabe S, Kakigi R, Miki K, Puce A. 2006. Human MT/V5 activity on viewing eye gaze changes in others: a magnetoencephalographic study. Brain Res. 1092:152–160.

Supplementary Figures

Supplementary Figure 1. **A**: Coronal sections of the anatomical mask for the right STS region. **B**: Saggital section showing the full length of the mask in the y plane.

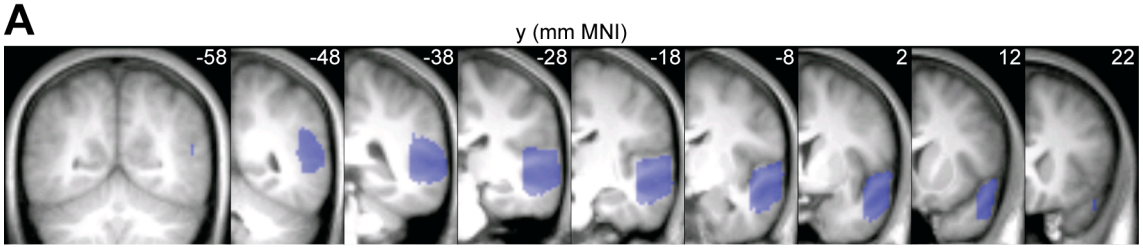
Supplementary Figure 2. Single volunteer searchlight results for the left-right head turn classification effect. Individual searchlight maps are thresholded at $p < 0.001$ (uncorrected, binomial test), and are masked to include only searchlight centers falling inside the right STS anatomical mask (highlighted). Codes for volunteers who were used in the final analysis of the follow-up scans are shown in the relevant panels. Note that single subject results were normalized and smoothed (10 mm FWHM) prior to the group analyses reported in the main text (see Materials and Methods), and thus any direct comparison between group and single volunteer results should be made with caution.

Supplementary Figure 3. Group results for the left STS anatomical ROI (highlighted, $p < 0.05$ FWE corrected within this region), displayed on the mean T1 volume for the sample. **A**: Classification of left-right ellipsoid rotation. **B**: Regions where left-right classification of ellipsoid rotations was more accurate than classification of head turns. **C**: Region with greater univariate responses to heads than to ellipsoids. **D**: Regions with greater univariate responses to ellipsoids than to heads.

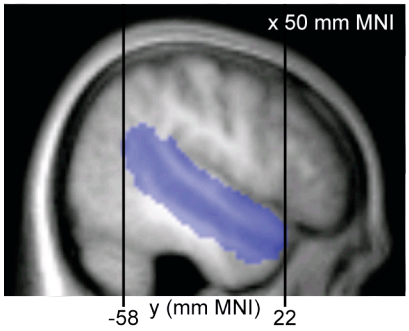
Supplementary Figure 4. Univariate responses to head turn direction in the original and revised experiments. **A**: Regions responding more to left than to right head turns (warm colors), and regions responding more to right than to left head turns (cool colors). Results are shown for the original experiment (without fixation) and for the revised experiment (with fixation) in rows, with the 3 volunteers from the final analysis of the revised experiment in

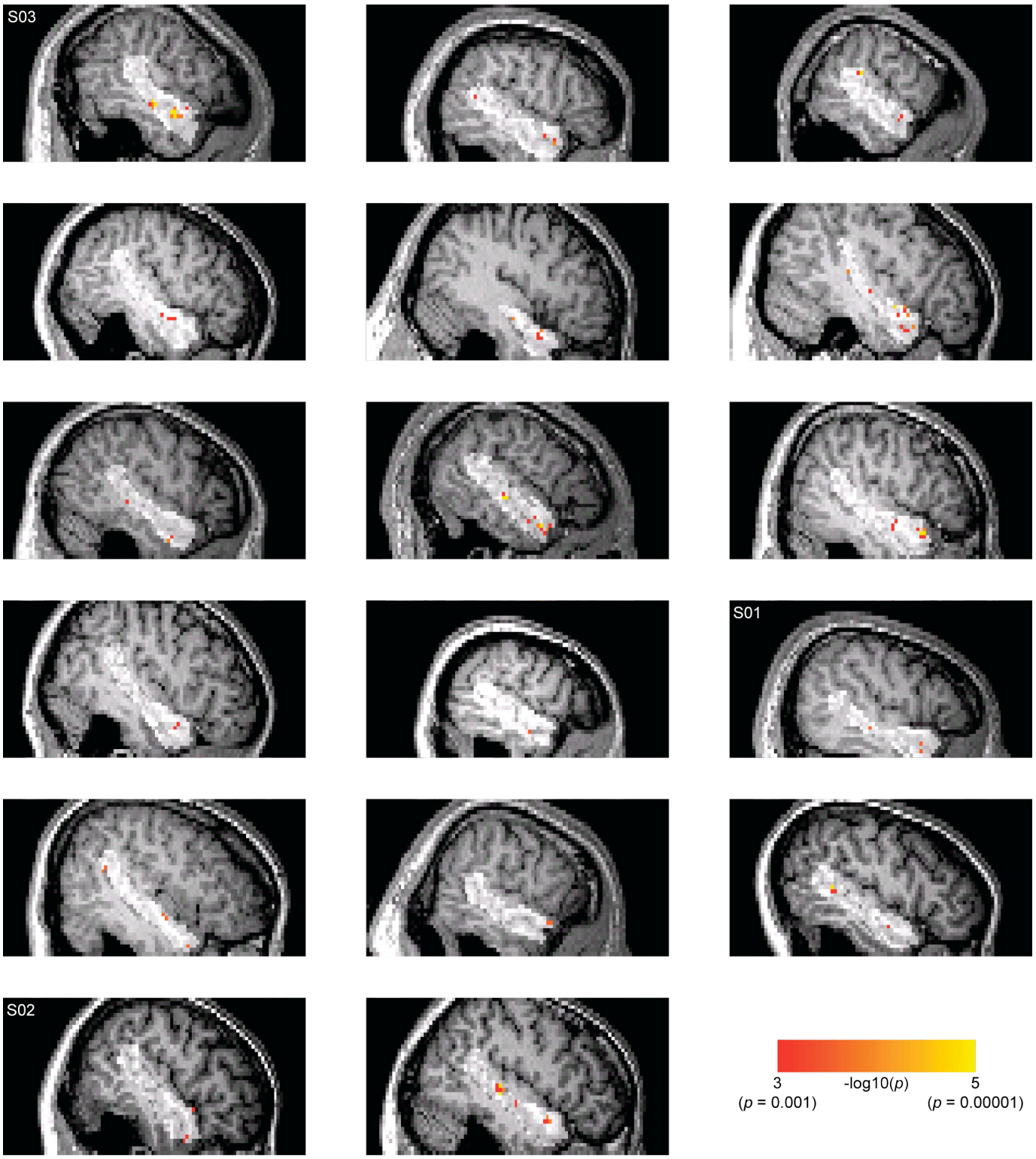
columns (S01, S02, S03). It can be seen that in the original experiment (without fixation), all selective responses in early visual cortex are ipsilateral to the direction of motion in the head stimuli (e.g. right early visual cortex responds more to rightward than to leftward head turns). In the follow-up experiment (with fixation), any selective responses in early visual cortex are instead contralateral to the direction of motion in the head stimuli. **B-C**: Selectivity for left against right head turns (*Z* statistics), plotted for the peak voxel in left and right early visual cortex, as selected by the contrast heads and ellipsoids against baseline. It can be seen that for all volunteers and hemispheres, the left-right head turn selectivity tends to become more contralateral in the revised experiment (with fixation), as compared to the original experiment (without fixation).

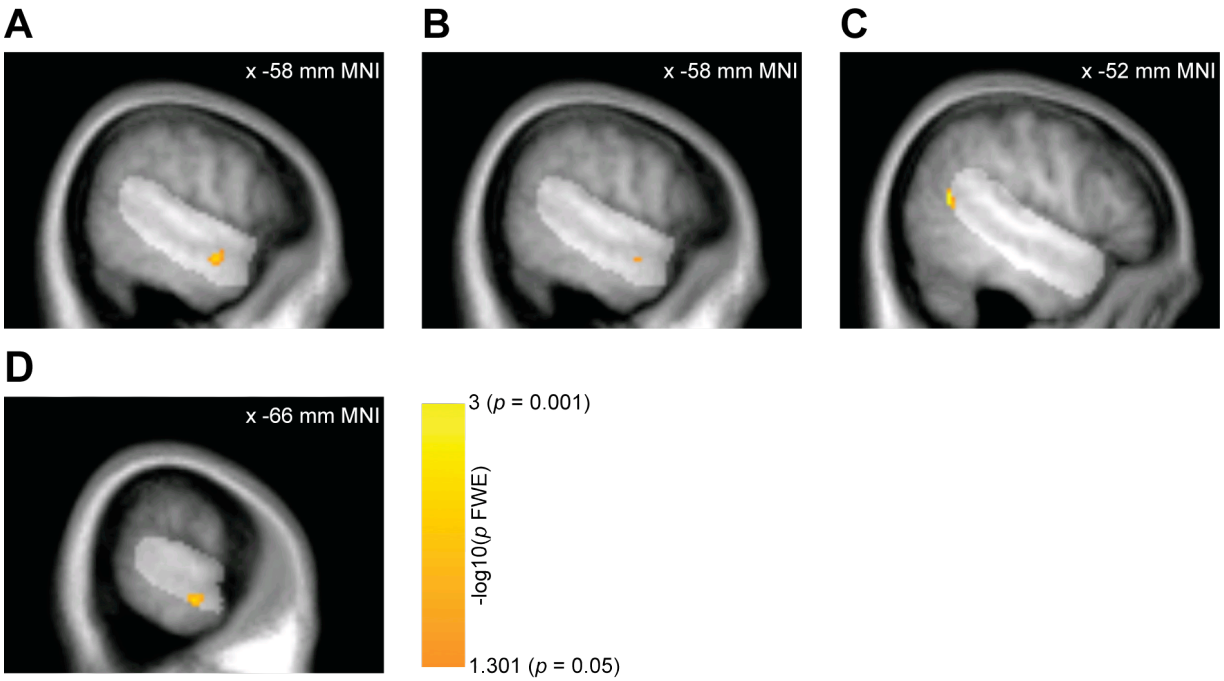
A

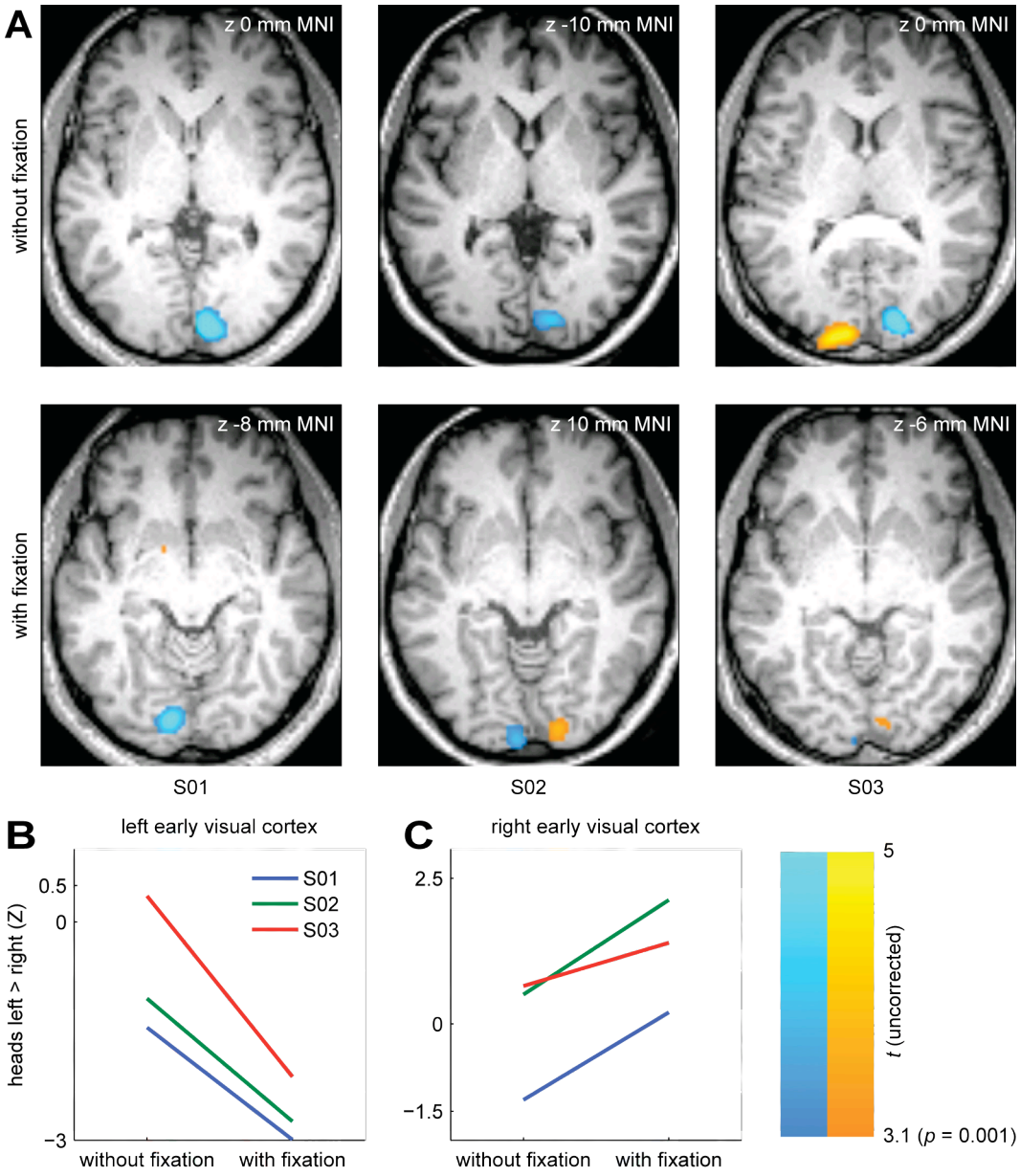


B




S03
S01
S02
3
-log10(p)
5
(p = 0.001)
(p = 0.00001)


A
x -58 mm MNI
B
x -58 mm MNI
C
x -52 mm MNI
D
x -66 mm MNI
3 (p = 0.001)
-log10(p FWE)
1.301 (p = 0.05)


A
without fixation
with fixation
z 0 mm MNI
z -10 mm MNI
z 0 mm MNI
z -8 mm MNI
z 10 mm MNI
z -6 mm MNI
S01
S02
S03
B
left early visual cortex
heads left > right (Z)
0.5
0
−3
S01
S02
S03
without fixation
with fixation
C
right early visual cortex
2.5
0
−1.5
without fixation
with fixation
5
t (uncorrected)
3.1 (p = 0.001)

Supplementary Tables

Supplementary Table 1. Regions with significant response level differences between heads and ellipsoids in the univariate analysis (p <0.05, FWE corrected for gray matter regions in the full volume).

Supplementary Table 2. Horizontal fixation change, analyzed using volunteer-specific ANOVAs with the factors of stimulus type (head, ellipsoid) and direction (leftward, rightward).

Supplementary Table 1. Regions with significant response level differences between heads and ellipsoids in the univariate analysis ($p < 0.05$, FWE corrected for gray matter regions in the full volume).

Region	Hemisphere	*p*	Peak (mm MNI)		
			x	*y*	*z*
Heads > Ellipsoids					
Early visual cortex	L/R	<0.001	18	-96	-4
Fusiform gyrus	R	<0.001	40	-44	-20
Middle temporal gyrus (MT+)	L	<0.001	-50	-72	14
Superior temporal sulcus	R	0.001	48	-44	16
Middle temporal gyrus (MT+)	R	0.001	52	-74	2
Middle frontal gyrus	R	0.016	46	4	52
Ellipsoids > Heads					
Parahippocampal gyrus	L	<0.001	-26	-52	-16
Parahippocampal gyrus	R	<0.001	28	-42	-12
Middle occipital gyrus	L	<0.001	-32	-90	10
Middle occipital gyrus	R	<0.001	36	-84	6
Anterior cingulate	L	0.007	-8	46	2
Lateral sulcus	R	0.007	58	2	2

Supplementary Table 2. Horizontal fixation change, analyzed using volunteer-specific ANOVAs with the factors of stimulus type (head, ellipsoid), and direction (leftward, rightward).

		F		
Volunteer	df	Stimulus type	Direction	Interaction
S01	1, 417	11.79*	6.48*	10.14*
S02	1, 382	2.29	17.92*	21.11*
S03	1, 418	8.36*	26.8*	31.75*
S04	1, 408	0.13	14.54*	51.16*
S05	1, 390	1.39	34.09*	76.72*

* $p < 0.05$

Appendix B: A head view-invariant representation of gaze direction in anterior superior temporal sulcus

A reprint of Experiments 5-6 (Chapter 4) in published form.

Please cite this article in press as: Carlin et al., A Head View-Invariant Representation of Gaze Direction in Anterior Superior Temporal Sulcus, Current Biology (2011), doi:10.1016/j.cub.2011.09.025

Current Biology *21*, 1–5, November 8, 2011 ©2011 Elsevier Ltd All rights reserved DOI 10.1016/j.cub.2011.09.025

Report

A Head View-Invariant Representation of Gaze Direction in Anterior Superior Temporal Sulcus

Johan D. Carlin,[1,*] Andrew J. Calder,[1] Nikolaus Kriegeskorte,[1] Hamed Nili,[1] and James B. Rowe[1,2,3]
[1]Medical Research Council Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 7EF, UK
[2]Department of Clinical Neurosciences, University of Cambridge, Addenbrooke's Hospital, Cambridge CB2 2QQ, UK
[3]Behavioural and Clinical Neuroscience Institute, University of Cambridge, Cambridge CB2 3EB, UK

Summary

Humans show a remarkable ability to discriminate others' gaze direction, even though a given direction can be conveyed by many physically dissimilar configurations of different eye positions and head views. For example, eye contact can be signaled by a rightward glance in a left-turned head or by direct gaze in a front-facing head. Such acute gaze discrimination implies considerable perceptual invariance. Previous human research found that superior temporal sulcus (STS) responds preferentially to gaze shifts [1], but the underlying representation that supports such general responsiveness remains poorly understood. Using multivariate pattern analysis (MVPA) of human functional magnetic resonance imaging (fMRI) data, we tested whether STS contains a higher-order, head view-invariant code for gaze direction. The results revealed a finely graded gaze direction code in right anterior STS that was invariant to head view and physical image features. Further analyses revealed similar gaze effects in left anterior STS and precuneus. Our results suggest that anterior STS codes the direction of another's attention regardless of how this information is conveyed and demonstrate how high-level face areas carry out fine-grained, perceptually relevant discrimination through invariance to other face features.

Results and Discussion

We designed a set of 25 computer-generated faces where nine gaze directions were conveyed by multiple, physically dissimilar configurations of different head views and eye positions (Figure 1A). This allowed us to disentangle functional magnetic resonance imaging (fMRI) responses consistent with head view-invariant representations of gaze direction from responses related to the faces' other physical features [2]. Previous reports of superior temporal sulcus (STS) involvement in perception of gaze and head view used faces in which eye position or head view were manipulated in isolation [3–5]. Such designs cannot address the issue of view-invariant coding of gaze because the degree of eye position or head view change defines the degree of gaze direction change. Moreover, previous attempts to identify view-invariant gaze codes using conventional univariate analysis of smoothed fMRI data have produced inconsistent results and did not observe gaze effects in STS [6, 7]. This is perhaps unsurprising, because macaque STS neurons that are selective for head view and gaze direction are organized into small patches [8, 9] beyond the likely resolution of conventional fMRI analysis methods. Recently, multivariate pattern analysis (MVPA) has been used to identify other visual representations thought to be coded at similarly small spatial scales, including direction-specific motion responses in early visual cortex [10, 11]. Here, we applied novel MVPA methods (representational similarity analysis [12]) to high-resolution fMRI data in order to reveal response pattern codes for view-invariant gaze direction.

Representational Similarity Analysis of Gaze Codes

Eighteen human participants carried out a one-back matching task while viewing the gazing faces in a rapid event-related fMRI experiment (for details, see Figure 1, Figure 2, and Experimental Procedures). Eye tracking data were also acquired to rule out confounding influences of eye movements (see Supplemental Experimental Procedures available online).

We extracted each participant's responses to each face (t contrast maps against baseline) to estimate response pattern dissimilarities between each face pair (1-Pearson r across voxels). These dissimilarities were compared to a predicted dissimilarity structure for view-invariant gaze direction and to other dissimilarity structures representing alternative accounts of the data (Figure 1B). We quantified the relationship between the response pattern dissimilarities and the predicted dissimilarities as the Spearman rank correlation across all face pairs. This representational similarity analysis [12] was carried out in single participants using a searchlight algorithm [13] (5 mm radius sphere) that localizes response pattern effects to local voxel neighborhoods.

Individual participants' results for each response-predictor comparison were spatially normalized to a common template, smoothed, and tested for group effects using a permutation test (Experimental Procedures). Based on previous evidence for right-lateralized gaze responses in human STS [3, 4, 14], we report all p values in the primary analysis corrected for multiple comparisons within the anatomically defined right STS region ($p < 0.05$, familywise error [FWE]; Figure S1A, 4598 voxels). For completeness, we also carried out exploratory analyses of left STS and the full gray-matter-masked volume.

Right STS Gaze Codes Are Invariant to Head View and Physical Stimulus Features

Response patterns in right anterior ($p = 0.013$) and posterior ($p = 0.006$) STS showed a consistent relationship with the view-invariant gaze direction predictor (Figure 3A). Complementary functional region of interest analyses of right STS revealed moderate independently estimated effect sizes in these regions ($r = 0.39$ for anterior STS, $r = 0.42$ for posterior; Figure S1B). Although these effects suggest that both regions code the direction of another's gaze, it was important to correct for unavoidable correlations between the view-invariant gaze direction predictor and alternative predictors derived from the faces' physical stimulus features (1-r across

*Correspondence: johan.carlin@mrc-cbu.cam.ac.uk

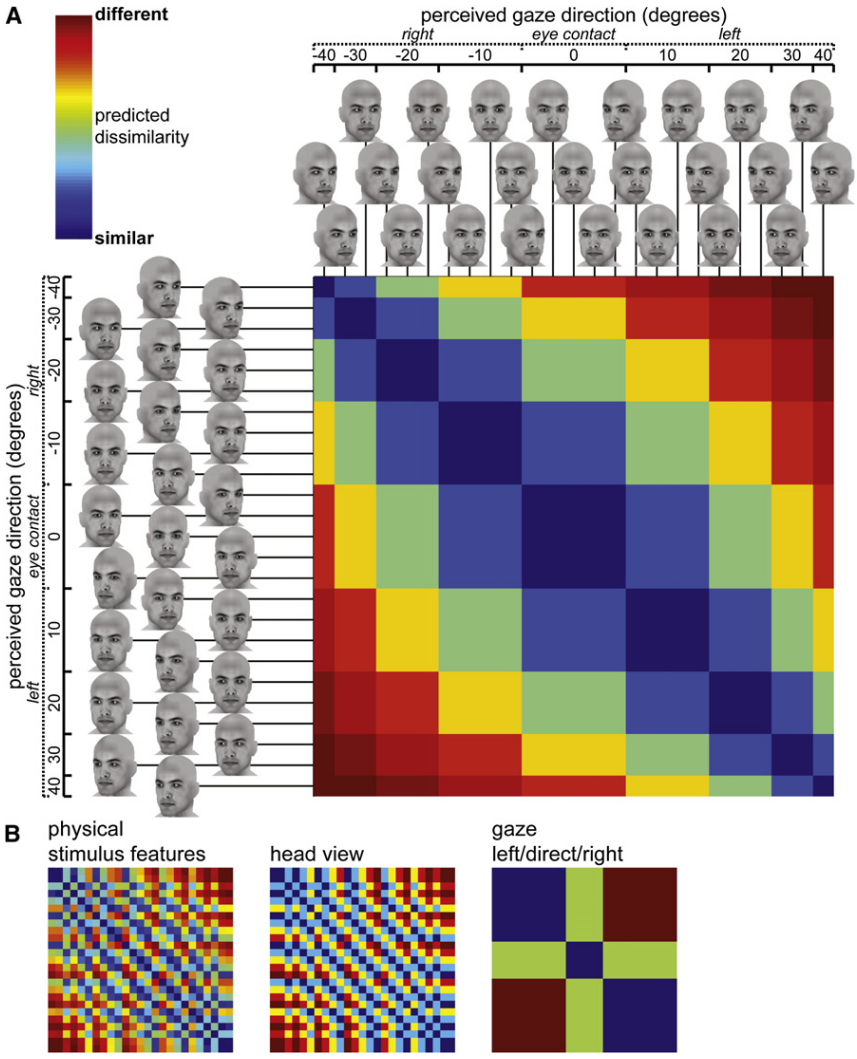


Figure 1. Stimuli and Predicted Dissimilarity Matrices

(A) Predicted view-invariant gaze direction dissimilarity structure across the 25 computer-generated faces. The faces are sorted according to the nine distinct gaze directions in the stimulus set (left 40° to right 40° rotation), which were created by incrementally varying head view and eye position relative to the head (five increments between left 20° and right 20° for both).

(B) Predicted dissimilarity structures for the same faces based on alternative accounts of the data corresponding to their physical stimulus features (1-r across image grayscale intensities), head view, and qualitative gaze direction (left/direct/right gaze) ignoring quantitative differences between angles of left and right gaze. Dissimilarity matrices are sorted as in (A).

image grayscale intensities) or head view (correlation between gaze direction and physical stimulus features r = 0.37, correlation between gaze direction and head view r = 0.36; Figure 1B). Note that the relationship with both is because the faces' physical stimulus features were almost entirely explained by head view (r = 0.99).

To exclude the contribution of these additional facial properties, we computed a further correlation between the view-invariant gaze direction predictor and the response pattern dissimilarities, this time partialing out any correlation between physical stimulus features and the response pattern dissimilarities (partial Spearman correlation). Only the perceived gaze direction effect in anterior STS remained significant when the influence of physical stimulus features was removed (p = 0.018; Figure 3B). Similarly, removing the influence of head view did not disrupt the effect of the view-invariant gaze direction predictor in anterior STS (p = 0.016) but produced only a weakly significant effect in posterior STS (p = 0.045; Figure 3C). Indeed, posterior STS showed a significant relationship with the predictor derived from the faces' physical stimulus features (p = 0.048) and a near-significant relationship with the predictor derived from head view (p = 0.08). Thus, gaze direction responses in posterior STS were influenced by physical stimulus features, which corresponded largely to variation in head view, whereas gaze direction responses in anterior STS were invariant to these facial properties.

Right STS Gaze Codes Are Fine Grained

If gaze codes in STS play a role in supporting perceptual performance, such codes should mirror human sensitivity to fine-grained gaze direction distinctions [15]. We tested this by comparing the original view-invariant gaze predictor representing nine gaze directions to a left/direct/right gaze predictor that distinguished between three qualitative gaze directions, while ignoring continuous information about the degree to which gaze is averted left or right (Figure 1B). Partial correlation analysis showed that the effects of the original view-invariant gaze predictor remained after removing the influence of the left/direct/right gaze predictor (anterior STS p = 0.016, posterior STS p = 0.018; Figure S1C). Thus, the reported view-invariant gaze direction effects cannot be explained by simpler gaze representations. Instead, gaze direction codes in STS contained fine-grained information about both the direction and the degree to which gaze is averted.

Gaze Codes in Left STS and Precuneus

An exploratory analysis of left STS revealed similar evidence of view-invariant coding of gaze direction in left anterior STS (Table S1). There were no significant effects in left posterior STS (p > 0.19). View-invariant representations of gaze direction in anterior STS may therefore be bilateral.

A further analysis of the full gray-matter-masked volume also revealed view-invariant gaze direction codes in precuneus, which survived all control analyses reported above (Table S1). Precuneus and STS are monosynaptically connected in macaques [16], and precuneus has previously been implicated in head/gaze following [17] and in attentional orienting [18], which suggests that gaze codes here may reflect gaze-cued shifts in attention [19]. Eye tracking analyses suggested that participants were fixating well (Supplemental

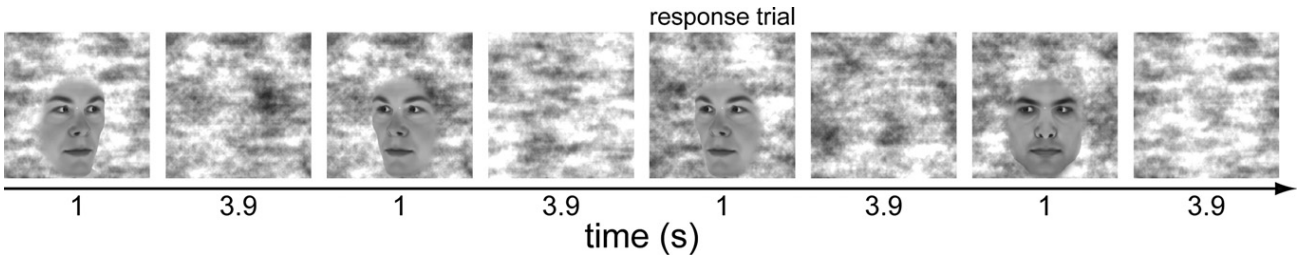


Figure 2. An Example Trial Sequence from the fMRI Experiment

The faces were presented in random order in a rapid event-related design. Participants maintained fixation on a central cross. The faces were presented so that the cross fell on the bridge of the nose of each face to minimize eye movements during the task. The 25 head/eye position configurations were posed by two identities (50 images total). Each was presented three times in five independently randomized sets (150 experimental trials presented over 11 min per set; 750 trials in total over 55 min). Each trial comprised a face (1 s) followed by an intertrial interval (2.9 s). Fifteen randomly selected trials in each set were immediately followed by a second presentation of the same face (75 added trials in total). Participants were asked to identify repetitions with a button response before the onset of the next trial (one-back task). Response trials were equally sampled from all head/eye position configurations and were modeled with a separate regressor of no interest in the first-level fMRI model. At the end of each set, participants viewed a feedback screen (20 s) that summarized their hit rates and false alarm rates for that set.

See Supplemental Experimental Procedures for a complete account of stimulus design and procedure.

Experimental Procedures), so these precuneus effects are likely driven by covert attentional shifts rather than overt eye movements.

Participant-Specific Gaze Codes

Our experimental design assumes that perceived gaze direction can be approximated by the sum angle of head view and eye position relative to the head (Figure 1) [2]. However, human gaze discrimination performance can be subtly biased by head view [20, 21]. We therefore carried out a follow-up behavioral experiment to assess whether the standard view-invariant gaze predictor we used was a good match for the participants' individual gaze discrimination performance. Each participant in the fMRI experiment carried out a subsequent task where they indicated the perceived gaze direction of the faces they had viewed in the scanner. Difference scores between the perceived gaze direction for the different face pairs were then compared to the standard view-invariant gaze predictor (Supplemental Experimental Procedures). Gaze discrimination performance was well explained by the generic view-invariant gaze direction predictor (median Spearman r = 0.90, 95% confidence = 0.87–0.93, bootstrap test), and this relationship survived removing the influence of each of the alternative predictors discussed above (Figure S1E).

We also repeated the fMRI analyses using the participant-specific gaze discrimination predictors in place of the standard view-invariant gaze direction predictor, and obtained comparable results (Table S1). Thus, participants' gaze

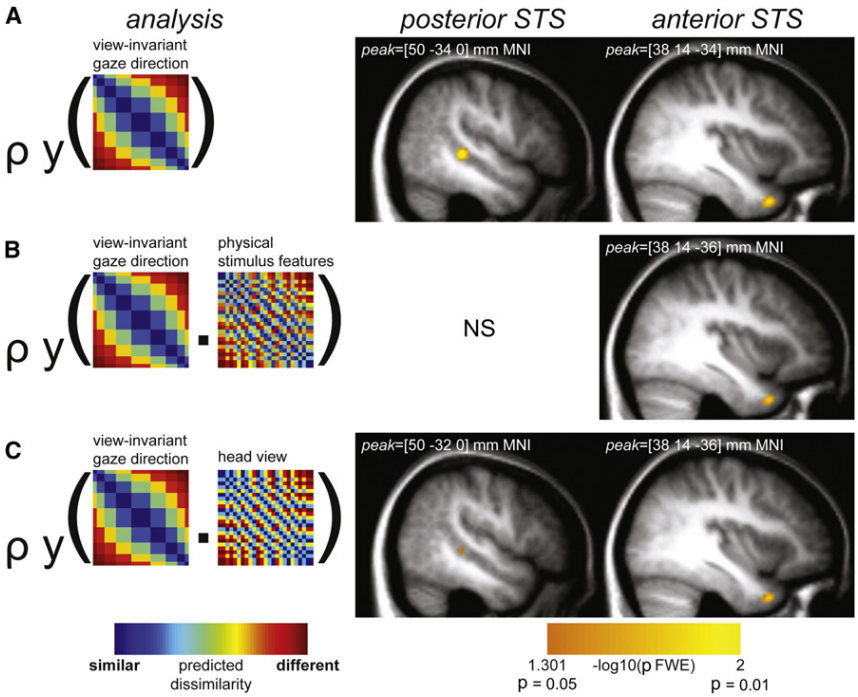


Figure 3. Regions with Pattern Responses to the Gazing Faces

Spearman correlations of partial correlation effects across participants (n = 18, p < 0.05, familywise error [FWE] corrected for right STS; Figure S1A) are shown overlaid on the sample's mean structural volume.

(A) Response pattern dissimilarities in anterior and posterior STS are explained by the view-invariant gaze direction predictor.

(B) Gaze direction responses in anterior STS alone are found for the same predictor when controlling for physical stimulus features.

(C) Similarly, gaze direction responses in anterior STS for the view-invariant gaze predictor are unaffected when controlling for head view, whereas responses in posterior STS are reduced.

discrimination performance was well approximated by the standard view-invariant gaze predictor, and the neural responses to the gazing faces were similarly explained by the standard and participant-specific gaze predictors.

Conclusions

This study provides the first evidence that human anterior STS contains a fine-grained, view-invariant code of perceived gaze direction. We also observed similar gaze effects in precuneus, which may reflect attentional orienting responses to gaze [19]. Our results do not rule out the existence of view-specific codes for particular head-gaze configurations but rather demonstrate that gaze perception is not achieved using such view-specific representations alone. Our results are consistent with the hypothesis that gaze perception is achieved through a high-level, view-invariant code of the direction of another's social attention in anterior STS.

The representational content of right posterior STS is distinct from anterior STS. Although the view-invariant gaze predictor also identified this region, this was largely accounted for by the modest correlation between this predictor and the faces' physical facial properties or head view, which showed significant or borderline relationships with the right posterior STS. This is consistent with recent work showing that response patterns in posterior STS can be used to distinguish head view [5]. The preferential involvement of anterior STS in view-invariant representations of gaze direction was further underlined by the analysis of left STS, which identified the anterior region alone. Our results are thus consistent with previous reports that right posterior STS is responsive to different gaze directions and head views [1, 5], but view-invariant gaze direction codes appear most prevalent in anterior STS.

Collectively, our results suggest a hierarchical processing stream for gaze perception, with increasing invariance to gaze-irrelevant features from posterior to anterior STS. Such a processing hierarchy would be consistent with recent evidence from neurons responsive to face identity in the macaque temporal lobe [22], where invariance to head view increases from middle STS to anterior inferotemporal cortex. Similarly, neurons tuned to specific head views in anterior STS also frequently respond to gaze direction [23–25], whereas neurons with head view tunings in middle STS generally do not [25]. Such hierarchical progressions toward view invariance may therefore be a general property of high-level face representations, regardless of whether these hierarchies serve to extract face identity or the direction of another's gaze.

In conclusion, response patterns in human anterior STS are not coded according to any readily observable visual face features but rather according to the direction of another person's gaze, irrespective of head view.

Experimental Procedures

Participants

Twenty-three right-handed participants with normal or corrected-to-normal vision were recruited for the study. Participants provided informed consent as part of a protocol approved by the Cambridge Psychology Research Ethics Committee. Five participants were removed from further analysis: two failed to complete the experiment, two fell asleep and displayed excessive head motion, and one failed to maintain fixation (Supplemental Experimental Procedures). This left 18 participants (five male, mean age 24, age range 18–36).

Imaging Acquisition

Scanning was carried out at the MRC Cognition and Brain Sciences Unit (Cambridge) using a 3 T TIM Trio Magnetic Resonance Imaging scanner (Siemens), with a head coil gradient set. Functional data were collected using high-resolution echo planar T2*-weighted imaging (40 oblique axial slices, repetition time [TR] 2490 ms, echo time [TE] 30 ms, in-plane resolution 2 × 2 mm, slice thickness 2 mm plus a 25% slice gap, 192 × 192 mm field of view). The acquisition window was tilted up approximately 30° from the horizontal plane to provide complete coverage of the occipital and temporal lobes. All volumes were collected in a single, continuous run for each participant. The initial six volumes from the run were discarded to allow for T1 equilibration effects. T1-weighted structural images were also acquired (MPRAGE, 1 mm isotropic voxels).

Imaging Analysis

Preprocessing of the fMRI data was carried out using Statistical Parametric Mapping 5 (SPM5; http://www.fil.ion.ucl.ac.uk/spm/). Structural volumes were segmented into gray- and white-matter partitions and normalized to the Montreal Neurological Institute (MNI) template using combined segmentation and normalization routines. All functional volumes were realigned to the first nondiscarded volume, slice time corrected, and coregistered to the T1 structural volume. The functional volumes remained unsmoothed and in their native space for participant-specific generalized linear modeling. Each set was modeled with a separate set of regressors for each head/eye configuration (25, collapsing across the two face identities), false alarms, and repeat trials. We also included scan nulling regressors to eliminate the effects of excessively noisy volumes [26, 27]. The experimental predictors were convolved with a canonical hemodynamic response function, and contrast images for each individual condition against the implicit baseline were generated based on the fitted responses. The resulting T contrast volumes were gray-matter-masked using the tissue probability maps generated by the segmentation processing stage and were used as inputs for representational similarity analysis.

Representational similarity analyses were carried out using custom code developed using Python and PyMVPA [28]. The voxels within each searchlight and each set were correlated across conditions (1-Pearson *r*), and the resulting 1-correlation matrix was averaged across the five sets to produce a final response pattern dissimilarity matrix for that searchlight. The data dissimilarities were then compared to a set of hypothesis-based predictors using the Spearman rank correlation or partial Spearman rank correlation. In all cases, the resulting correlation coefficient was Fisher transformed and mapped back to the central voxel in the searchlight, yielding a descriptive individual subject map that was entered into a group analysis. This two-stage summary statistics procedure resembles that used in conventional univariate fMRI group analysis [29]. The individual subject maps were normalized to the MNI template and were smoothed to overcome errors in intersubject alignment (10 mm full width at half mean [FWHM]). The resulting volumes were entered into a permutation-based random-effects analysis using statistical nonparametric mapping [30] (SnPM; 10,000 permutations, 10 mm FWHM variance smoothing). The use of nonparametric tests avoids distributional assumptions regarding the nature of the descriptive maps and avoids inherent problems in applying standard SPM5 FWE correction based on random Gaussian fields to discontinuous gray-matter-masked data.

Supplemental Information

Supplemental Information includes one figure, one table, and Supplemental Experimental Procedures and can be found with this article online at doi:10.1016/j.cub.2011.09.025.

Acknowledgments

We are grateful to Raliza Stoyanova and Doris Tsao for helpful comments on previous versions of this manuscript and to Ian Nimmo-Smith for advice on statistics. This work was supported by the United Kingdom Medical Research Council (MC_US_A060_0017 to A.J.C.; MC_US_A060_0016 to J.B.R.) and the Wellcome Trust (WT088324 to J.B.R.).

Received: July 14, 2011
Revised: September 13, 2011
Accepted: September 13, 2011
Published online: October 27, 2011

References

1. Nummenmaa, L., and Calder, A.J. (2009). Neural mechanisms of social attention. Trends Cogn. Sci. (Regul. Ed.) *13*, 135–143.

2. Todorović, D. (2006). Geometrical basis of perception of gaze direction. Vision Res. *46*, 3549–3562.
3. Calder, A.J., Beaver, J.D., Winston, J.S., Dolan, R.J., Jenkins, R., Eger, E., and Henson, R.N.A. (2007). Separate coding of different gaze directions in the superior temporal sulcus and inferior parietal lobule. Curr. Biol. *17*, 20–25.
4. Carlin, J.D., Rowe, J.B., Kriegeskorte, N., Thompson, R., and Calder, A.J. (2011). Direction-sensitive codes for observed head turns in human superior temporal sulcus. Cereb. Cortex. Published online June 27, 2011. 10.1093/cercor/bhr061.
5. Natu, V.S., Jiang, F., Narvekar, A., Keshvari, S., Blanz, V., and O'Toole, A.J. (2010). Dissociable neural patterns of facial identity across changes in viewpoint. J. Cogn. Neurosci. *22*, 1570–1582.
6. Pageler, N.M., Menon, V., Merin, N.M., Eliez, S., Brown, W.E., and Reiss, A.L. (2003). Effect of head orientation on gaze processing in fusiform gyrus and superior temporal sulcus. Neuroimage *20*, 318–329.
7. George, N., Driver, J., and Dolan, R.J. (2001). Seen gaze-direction modulates fusiform activity and its coupling with other brain areas during face processing. Neuroimage *13*, 1102–1112.
8. Perrett, D.I., Smith, P.A., Potter, D.D., Mistlin, A.J., Head, A.S., Milner, A.D., and Jeeves, M.A. (1984). Neurones responsive to faces in the temporal cortex: studies of functional organization, sensitivity to identity and relation to perception. Hum. Neurobiol. *3*, 197–208.
9. Wang, G., Tanifuji, M., and Tanaka, K. (1998). Functional architecture in monkey inferotemporal cortex revealed by in vivo optical imaging. Neurosci. Res. *32*, 33–46.
10. Kamitani, Y., and Tong, F. (2006). Decoding seen and attended motion directions from activity in the human visual cortex. Curr. Biol. *16*, 1096–1102.
11. Seymour, K., Clifford, C.W., Logothetis, N.K., and Bartels, A. (2009). The coding of color, motion, and their conjunction in the human visual cortex. Curr. Biol. *19*, 177–183.
12. Kriegeskorte, N., Mur, M., and Bandettini, P. (2008). Representational similarity analysis - connecting the branches of systems neuroscience. Front. Sys. Neurosci. *2*, 4.
13. Kriegeskorte, N., Goebel, R., and Bandettini, P.A. (2006). Information-based functional brain mapping. Proc. Natl. Acad. Sci. USA *103*, 3863–3868.
14. Pelphrey, K.A., Singerman, J.D., Allison, T., and McCarthy, G. (2003). Brain activation evoked by perception of gaze shifts: the influence of context. Neuropsychologia *41*, 156–170.
15. Symons, L.A., Lee, K., Cedrone, C.C., and Nishimura, M. (2004). What are you looking at? Acuity for triadic eye gaze. J. Gen. Psychol. *131*, 451–469.
16. Seltzer, B., and Pandya, D.N. (1994). Parietal, temporal, and occipital projections to cortex of the superior temporal sulcus in the rhesus monkey: a retrograde tracer study. J. Comp. Neurol. *343*, 445–463.
17. Laube, I., Kamphuis, S., Dicke, P.W., and Thier, P. (2011). Cortical processing of head- and eye-gaze cues guiding joint social attention. Neuroimage *54*, 1643–1653.
18. Cavanna, A.E., and Trimble, M.R. (2006). The precuneus: a review of its functional anatomy and behavioural correlates. Brain *129*, 564–583.
19. Friesen, C., and Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. Psychon. Bull. Rev. *5*, 490–495.
20. Gibson, J.J., and Pick, A.D. (1963). Perception of another person's looking behavior. Am. J. Psychol. *76*, 386–394.
21. Gamer, M., and Hecht, H. (2007). Are you looking at me? Measuring the cone of gaze. J. Exp. Psychol. Hum. Percept. Perform. *33*, 705–715.
22. Freiwald, W.A., and Tsao, D.Y. (2010). Functional compartmentalization and viewpoint generalization within the macaque face-processing system. Science *330*, 845–851.
23. Perrett, D.I., Hietanen, J.K., Oram, M.W., and Benson, P.J. (1992). Organization and functions of cells responsive to faces in the temporal cortex. Philos. Trans. R. Soc. Lond. B Biol. Sci. *335*, 23–30.
24. Perrett, D.I., Smith, P.A., Potter, D.D., Mistlin, A.J., Head, A.S., Milner, A.D., and Jeeves, M.A. (1985). Visual cells in the temporal cortex sensitive to face view and gaze direction. Proc. R. Soc. Lond. B Biol. Sci. *223*, 293–317.
25. De Souza, W.C., Eifuku, S., Tamura, R., Nishijo, H., and Ono, T. (2005). Differential characteristics of face neuron responses within the anterior superior temporal sulcus of macaques. J. Neurophysiol. *94*, 1252–1266.
26. Lemieux, L., Salek-Haddadi, A., Lund, T.E., Laufs, H., and Carmichael, D. (2007). Modelling large motion events in fMRI studies of patients with epilepsy. Magn. Reson. Imaging *25*, 894–901.
27. Rowe, J.B., Eckstein, D., Braver, T., and Owen, A.M. (2008). How does reward expectation influence cognition in the human brain? J. Cogn. Neurosci. *20*, 1980–1992.
28. Hanke, M., Halchenko, Y., Sederberg, P., Olivetti, E., Fründ, I., Rieger, J., Herrmann, C., Haxby, J., Hanson, S.J., and Pollmann, S. (2009). PyMVPA: A unifying approach to the analysis of neuroscientific data. Front. Neuroinform. *3*, 3.
29. Holmes, A., and Friston, K. (1998). Generalisability, random effects & population inference. Neuroimage *7*, S754.
30. Nichols, T.E., and Holmes, A.P. (2002). Nonparametric permutation tests for functional neuroimaging: a primer with examples. Hum. Brain Mapp. *15*, 1–25.

Current Biology, Volume *21*

Supplemental Information

A Head View-Invariant Representation of Gaze Direction in Anterior Superior Temporal Sulcus

Johan D. Carlin, Andrew J. Calder, Nikolaus Kriegeskorte, Hamed Nili, and James B. Rowe

Supplemental Inventory

Figure S1. Supplemental results relating to Figure 3.

Table S1. Analyses of additional regions and tests of the performance of a participant-specific gaze discrimination predictor in place of the standard view-invariant gaze direction predictor (related to Figure 3).

Supplemental Experimental Procedures

Stimulus Design and Presentation. Description of the methods employed to generate the face images used in the experiment, and the exact parameters under which they were presented during the experiment (related to Figure 1).

Behavioral Performance. Description of control analysis of in-scanner behavioral task (related to Figure 1).

Eye Tracking. Description of acquisition and control analysis of in-scanner eye tracking data (related to Figure 1).

Gaze Discrimination Experiment. Description of control experiment to support our conception of gaze direction as the sum of head view and eye view relative to head view (related to Figure 2).

Supplemental References

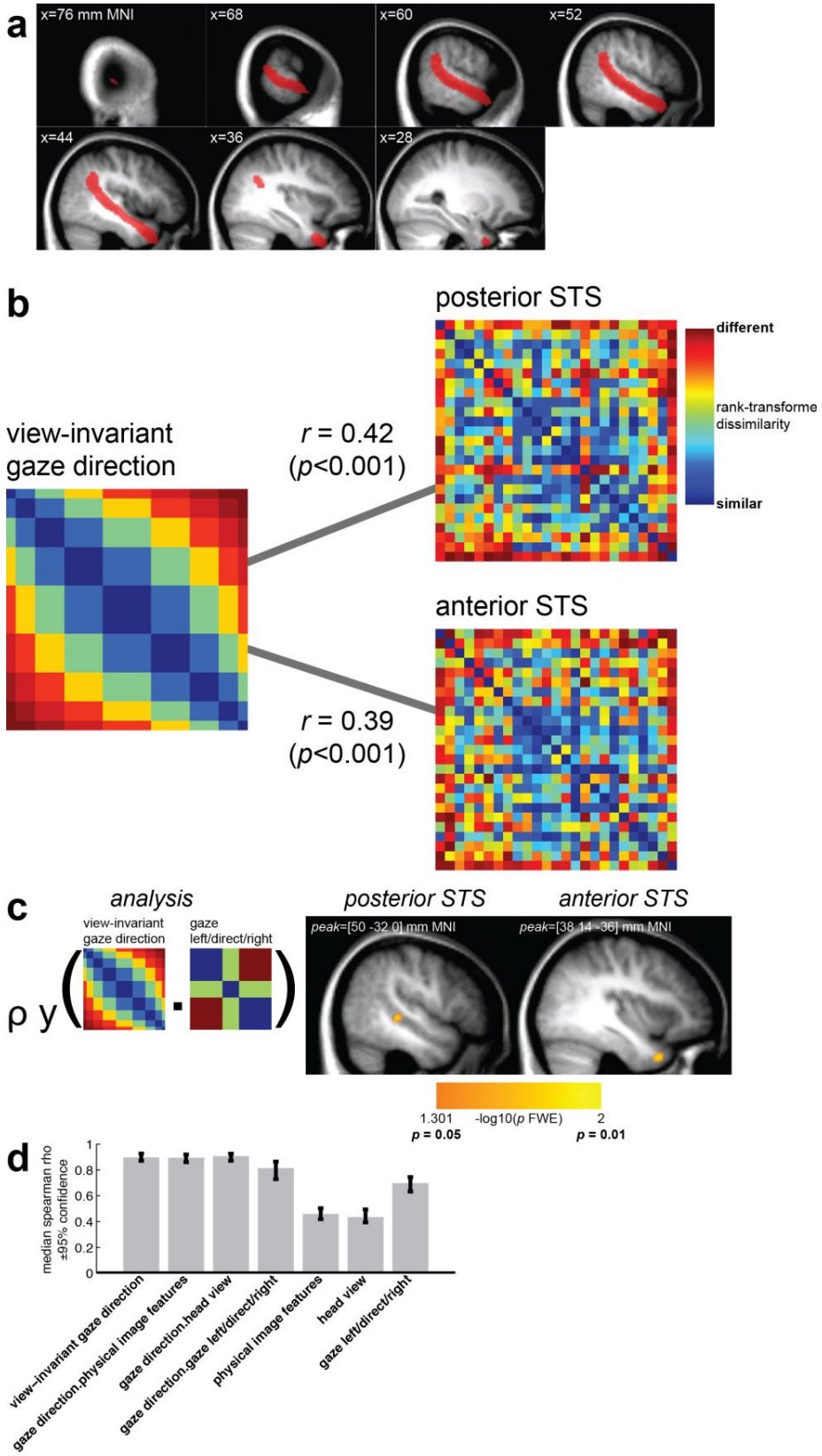

a
x=76 mm MNI
x=68
x=60
x=52
x=44
x=36
x=28
b
posterior STS
different
rank-transforme
dissimilarity
similar
view-invariant
gaze direction
r = 0.42
(p<0.001)
anterior STS
r = 0.39
(p<0.001)
c
analysis
view-invariant
gaze direction
gaze
left/direct/right
ρ y
posterior STS
anterior STS
peak=[50 -32 0] mm MNI
peak=[38 14 -36] mm MNI
1.301
-log10(p FWE)
2
p = 0.05
p = 0.01
d
median spearman rho
±95% confidence
1
0.8
0.6
0.4
0.2
0
view–invariant gaze direction
gaze direction.physical image features
gaze direction.head view
gaze direction.gaze left/direct/right
physical image features
head view
gaze left/direct/right

Figure S1.

Figure S1. Supplemental Results Relating to Figure 3

(A) Spatial extent of the right STS anatomical mask, shown overlaid on the sample's mean T1 volume.

(B) Independently estimated correlation between the view-invariant gaze direction predictor and response pattern dissimilarities in anterior and posterior right STS regions. Regions of interest (ROI) were defined using a leave-one-set-out procedure. We carried out a group analysis (similar parameters as main analysis) separately for the ROI-defining data in each unique split (4 of 5 sets) of the data to identify response pattern dissimilarities that were explained by view-invariant gaze direction. Responses to each set were estimated in five separate first-level models with 7 discarded volumes (17.43 s) separating each model to ensure independent estimates. Statistical thresholds for ROI definition varied between splits ($p < 0.01$ to $p < 0.05$, uncorrected). The only regions that appeared consistently across splits were anterior STS (mean [33.2, 10.0, -41.2] mm MNI, standard deviation [1.0, 5.1, 1.0]), and posterior STS (mean [46.4, -29.8, 4.0] mm MNI, standard deviation [3.2, 3.5, 4.2]). To better accommodate alignment errors across participants, we identified the participant-specific peak within a 10 mm radius of each group peak using ROI-defining data only. Subsequent tests of the identified ROIs were carried out separately for each split (e.g., ROIs defined using sets 1-4 were tested using set 5). We generated the illustrated response pattern dissimilarities for anterior and posterior STS by first averaging each participant's dissimilarities for each ROI across the 5 independent test splits, and then averaging the resulting ROI dissimilarity matrices across participants. It can be seen that both anterior and posterior STS showed consistent effects of view-invariant gaze direction in the independent test data (*p* values were defined using a permutation test where the order of the matrices were shuffled without replacement 10000 times [1]).

(C) Fine-grained gaze direction codes in right STS. Regions with consistent pattern responses (partial Spearman correlation) across participants ($n = 18$, $p < 0.05$ FWE). View-invariant gaze direction responses in anterior and posterior right STS remain when the influence of a qualitative distinction between gaze left/direct/right is removed.

(D) Gaze direction discrimination. Median Spearman correlations (bars 1,5-7) and median partial Spearman correlations (bars 2-4) across the participants (+/- 95% bootstrap confidence intervals). The participants' gaze discrimination performance was most strongly correlated with the view-invariant gaze direction predictor. Although performance was also moderately correlated with physical image features and head view, the strength of the relationship between discrimination performance and the view-invariant gaze direction predictor is relatively unaffected by partialling out the influence of these alternative predictors.

Table S1. Analyses of Additional Regions and Tests of the Performance of a Participant-Specific Gaze Discrimination Predictor in Place of the Standard View-Invariant Gaze Direction Predictor (Related to Figure 3)

Analysis	Comparison	Region	*p* (FWE)	Peak (mm MNI) *x*	*y*	*z*
right STS, 4598 voxels (p<0.05)	View-invariant gaze direction	posterior STS	0.01	50	-34	0
		anterior STS	0.01	38	14	-34
	… partialling out physical stimulus features	anterior STS	0.02	38	14	-36
	… partialling out head view	posterior STS	0.05	50	-32	0
		anterior STS	0.02	38	14	-36
	… partialling out gaze left/direct/right	posterior STS	0.02	50	-32	0
		anterior STS	0.02	38	14	-36
	Physical stimulus features	posterior STS	0.05	50	-32	2
	Participant-specific gaze discrimination	posterior STS	0.01	50	-34	0
		anterior STS	0.04	38	12	-36
	… partialling out physical stimulus features	anterior STS	0.05	38	12	-36
		posterior STS	0.06	50	-34	0
	… partialling out head view	posterior STS	0.04	50	-34	0
		anterior STS	0.05	38	12	-36
	… partialling out gaze left/direct/right	posterior STS	0.01	50	-34	0
		anterior STS	0.04	38	12	-36
left STS, 4442 voxels (p<0.1 to illustrate marginally significant effects mirroring those observed in right STS)	View-invariant gaze direction	anterior STS	0.04	-60	-8	-16
	… partialling out physical stimulus features	anterior STS	0.05	-62	-6	-16
	… partialling out head view	anterior STS	0.05	-62	-6	-16
	… partialling out gaze left/direct/right	anterior STS	0.03	-60	-8	-16
		middle STS	0.07	-66	-26	-8
	Participant-specific gaze discrimination	anterior STS	0.05	-60	-6	-16
	… partialling out physical stimulus features	anterior STS	0.07	-62	-6	-14
	… partialling out head view	anterior STS	0.05	-62	-6	-14
	… partialling out gaze left/direct/right	anterior STS	0.03	-60	-8	-16
		middle STS	0.06	-66	-26	-8
whole brain analysis, 134174 voxels (p<0.05)	View-invariant gaze direction	precuneus	0.01	4	-58	30
		cerebellum	0.04	-16	-76	-50
		cingulate gyrus	0.05	2	-22	32
	… partialling out physical stimulus features	precuneus	0.02	4	-58	30
	… partialling out head view	precuneus	0.01	4	-58	30
		cerebellum	0.05	-18	-76	-50
	… partialling out gaze left/direct/right	precuneus	0.03	4	-58	30
		cerebellum	0.02	-16	-76	-50
		occipital pole	0.04	-10	-98	2
	Participant-specific gaze discrimination	precuneus	0.01	4	-58	30
		posterior STS	0.04	50	-34	0
	… partialling out physical stimulus features	precuneus	0.03	4	-58	30
	… partialling out head view	precuneus	0.02	4	-58	30
		cerebellum	0.04	-18	-76	-50
	… partialling out gaze left/direct/right	precuneus	0.01	6	-58	30
		posterior STS	0.03	50	-34	0
		anterior STS	0.04	-60	-12	-26

Peak MNI coordinates are shown with p values FWE-corrected for regions as indicated by the analysis column.

Supplemental Experimental Procedures

Stimulus Design and Presentation

We used Poser 6 (Curious Labs Inc. Santa Cruz, CA) to create grey scale face images of two identities, each displaying 25 head-gaze configurations. Each face varied in horizontal head view (5 increments from left 20° to right 20°), horizontal eye position relative to the head (same increments as for head view), and identity (2 faces). The faces were processed in Matlab (The MathWorks, Inc., Natick, MA) to achieve similar luminance histograms, and were cropped to ensure that each face appeared in a similar retinal area. Cropping was achieved with a smooth border, and the resulting face was superimposed on a background texture that varied across conditions and across repetitions of the same face (Figure 1a). The background textures were created by Fourier-scrambling each of the 50 faces separately. The inclusion of the background texture served to reduce the influence of low-level physical differences between the conditions, and to increase the difficulty level of the one-back behavioral task. Stimuli were back-projected onto a screen in the scanner which participants viewed via a tilted mirror. The stimuli extended a 6° visual angle including the background texture, and approximately 3° horizontally by 4° vertically without it. The experiment was controlled using Matlab and the Psychophysics toolbox [2].

Behavioral Performance

Participants carried out a one-back face matching task whilst in the scanner. The task instruction was to respond to any repetition of the same face (same identity and head view/eye position configuration) while ignoring the scrambled backgrounds, which did not repeat. Accuracy was relatively high across the sample (mean 77%, standard error 3%), with low false alarm rates (mean 4.5% of trials, standard error 2.3%), and high sensitivity (mean d' 2.52, standard error 0.13). The large number of different head view/eye position configurations (25) relative to the number of response trials (75 per participant) meant that there was insufficient behavioral data available to model each of the 25 configurations separately. Thus, we pooled the available response trials according to gaze direction, and calculated accuracy scores for each of the 9 gaze directions. Repeated-measures ANOVA revealed no significant accuracy effects of gaze direction (p>0.23), suggesting that attentional or performance differences did not confound our fMRI analysis.

Eye Tracking

All participants' eye movements were monitored in the scanner using an infrared video-based eye tracker (50 Hz acquisition, Sensomotoric Instruments, Germany). Successful calibrations were obtained for 10 participants out of the final sample of 18. The remaining participants were excluded from the eye tracking analysis. On-line visual inspection of the eye tracking monitor suggested that these participants were maintaining their gaze at the fixation cross. Eye tracking data were analysed using custom code developed in Matlab.

To measure stimulus-induced eye movements, we analysed how the horizontal and vertical fixation position shifted between the start and the end of each stimulus presentation. This fixation shift was analysed using ANOVAs for each individual participant. We used a one-way ANOVA where the faces were labelled according to perceived gaze direction. One participant showed an effect of gaze direction on horizontal fixation shifts ($F_{(711,8)}$=2.37, p = 0.016). This participant was removed from further analyses of the fMRI data. No other horizontal or vertical

fixation shift effects were significant in single subjects, or in a group analysis. Thus, our participants were able to comply with the instruction to maintain fixation on the central cross.

Gaze Discrimination Experiment

At the end of the fMRI experiment, all participants carried out a follow-up gaze direction discrimination experiment outside the scanner. Participants viewed the stimuli from the main fMRI experiment from a fixed position in a head rest. On each trial, participants were asked to indicate the perceived direction of gaze by rotating a pointer in the horizontal plane; this pointer was positioned in front of the screen directly underneath the stimulus. Separate representational dissimilarity matrices were then generated from each participant's discrimination data by computing the difference between the mean pointer positions for each pair of head view/eye position configurations. The relationship between each participant's perceptual discrimination and the predicted dissimilarity structures was then estimated using a similar procedure to the main fMRI analysis. Inference was carried out at the group level, using confidence intervals based on bootstrap testing of the median (bootci function in Matlab, 95% interval, 10000 samples).

Participants' gaze discrimination performance was highly correlated with the view-invariant gaze direction predictor (Figure S1d), and to a lesser extent also with physical image features and head view. The correlation between discrimination performance and the view-invariant gaze direction predictor was largely unaffected when the influence of each of these alternative predictors was partialled out. The standard gaze direction predictor was thus a good approximation for participants' perceived gaze direction in this stimulus set.

Supplemental References

1. Kriegeskorte, N., Mur, M., and Bandettini, P. (2008). Representational similarity analysis - connecting the branches of systems neuroscience. Front. Sys. Neurosci. *2*, 1-28.
2. Brainard, D. (1997). The psychophysics toolbox. Spat. Vis. *10*, 433-436.