Pure

Bangor University

DOCTOR OF PHILOSOPHY

Body selectivity in human visual cortex

Peelen, Marius

Award date:
2006

Awarding institution:
Bangor University

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 17. Apr. 2024

University of Wales

Prifysgol Cymru

Body selectivity in human visual cortex

Marius Vincent Peelen

This thesis is submitted in part fulfilment of the degree of Doctor of Philosophy, completed at the Centre for Cognitive Neuroscience, School of Psychology, University of Wales, Bangor.

P.C.B. LLYFRGELL U.W.B LIBRARY

Acknowledgements

First of all, I would like to thank Paul Downing for his endless support and supervision. He taught me all aspects of cognitive neuroscience research, from experimental design to writing and submitting papers. He also introduced me to other facets of science by engaging me in the peer review process, student supervision, grant application and university politics. I could not have wished for anything more when I started my PhD. I would also like to thank the other members of my committee, Bob Rafal and Steve Tipper, for their valuable input during the bi-annual review meetings.

My fellow lab members Alison Wiggett, Annie Chan, and India Morrison have helped with much of the work in this thesis, thanks! I was very lucky with the arrival of David Linden, who has read and commented on most of the chapters in this thesis, which has certainly facilitated the subsequent publication of these chapters. The members of the Bangor imaging group have provided a very stimulating research environment, with discussions ranging from practical issues such as consent forms and cortex-based alignment to more theoretical issues like oscillatory synchronization and the neural basis of the BOLD signal. Tony Bedson has been invaluable for his help in data collection and data transfer. After busy days in the hospital he spent many nights helping out with testing participants, while at the same time sharing his knowledge about magnetic resonance imaging and the diseases it can diagnose.

My Bangor friends: Berenice, Chris, Claire, Corina, Dean, India, Jason, Jorien, Julia, Martijn, Paola, Patric, Richard, Ruth, Werner; thank you for stimulating discussions after work, and for making me feel at home in Bangor.

Finally, I would like to thank my parents and my brother and sister for their emotional support and their belief in me.

Parts of this work have been communicated as the following:

Chapter 2:
Peelen MV, Downing PE. (2005) Within-subject reproducibility of category-specific visual activation with functional MRI. *Human Brain Mapping 25*:402-408

Chapter 3:
Peelen MV, Downing PE. (2005) Selectivity for the human body in the fusiform gyrus. *Journal of Neurophysiology 93*:603-608

Chapter 4:
Peelen MV, Wiggett AJ, Downing PE. (2006) Patterns of fMRI activity dissociate overlapping functional brain areas that respond to biological motion. *Neuron 49*:815-822

Chapter 5:
Downing PE, Peelen MV, Wiggett AJ, Tew BD. (2006) The role of the extrastriate body area in action perception. *Social Neuroscience 1*:52-62

Chapter 6:
Peelen MV, Downing PE. (2005) Is the extrastriate body area involved in motor actions? *Nature Neuroscience 8*:125

CONTENTS

Summary 1

Chapter 1 –Introduction 2

Functional magnetic resonance imaging 4

The visual system 7

Overview of thesis 21

Chapter 2 – Within-subject reproducibility of category-specific visual activation with functional MRI 25

Abstract 26

Introduction 27

Materials and Methods 30

Results 36

Discussion 41

Chapter 3 – Selectivity for the human body in the fusiform gyrus 44

Abstract 45

Introduction 46

Materials and Methods 49

Results 54

Discussion 60

Chapter 4 – Patterns of fMRI activity dissociate overlapping functional brain areas that respond to biological motion 64

Abstract 65

Introduction 66

Materials and Methods 68

Results 74

Discussion 89

Chapter 5 – The role of the extrastriate body area in action perception 96

Abstract 97

Introduction 98

Materials and Methods 104

Results 108

Discussion 114

Chapter 6 – Is the extrastriate body involved in motor actions? 119

Chapter 7 – General Discussion 124

Organization of the posterior fusiform gyrus 127

The role of the EBA in body perception 132

Conclusions and future directions 139

References 142

SUMMARY

Perceiving other people is a seemingly effortless process. Yet within a few hundred milliseconds we are aware of who we are looking at, what this person is doing, and even what this person feels. We derive this information from the form and motion of the face and body. Faces may be particularly important for some aspects of person perception (e.g., identity recognition), whereas bodies may be more important for others (e.g., action recognition). Furthermore, information from the body is important in cases where it is not possible to perceive the details of the face, for instance when the face is occluded, or when we see someone from a distance. In most cases, however, it is likely that information from both the face and the body are perceived in parallel and are integrated at an early stage. Previous research on person perception has mostly focused on the brain mechanisms underlying face perception. Much less research has focused on the brain mechanisms underlying body perception, which is the topic of this thesis.

Using functional magnetic resonance imaging (fMRI) I provide evidence for a previously unknown body-selective visual area that overlaps a face-selective area. By employing novel analysis techniques that take into account patterns of activation across voxels I show that body- and face-selective areas can be functionally dissociated. Finally, I show that, in contrast to frontal and parietal action-recognition areas, visual body-selective areas do not contain a dynamic representation of observed actions.

Together, these findings increase our understanding of the brain mechanisms underlying body, face and action perception, by showing both similarities and dissimilarities in the brain structures involved in these processes.

Chapter 1:

Introduction

This thesis investigates body-selective fMRI responses in human visual cortex. Much of the research described here follows from earlier work on category-selective visual responses in monkeys and humans (Allison et al., 1994; 1999; Desimone and Gross, 1979; Desimone et al., 1984; Downing et al., 2001a; Epstein and Kanwisher, 1998; Puce et al., 1995; 1996; Kanwisher et al., 1997; Perrett et al., 1982; 1985; Rolls, 1984; Wachsmuth et al., 1994). Of direct relevance to the current work is the paper by Downing and colleagues that describes a body-selective area in human extrastriate cortex (Downing et al., 2001a). The discovery of this body-selective area raised many questions regarding its functional properties and its interaction with other brain areas involved in perceiving people. In the five experimental chapters of this thesis I will try to address some of these questions.

In this chapter I will provide some background that may help in understanding the experimental chapters and the scientific context in which the research was conducted. First, I will give an overview of the analysis methods that I used, insofar they deviate from standard fMRI methods. Of particular importance is the use of ROI analysis and voxelwise pattern analysis. Then, I will briefly review research on the human visual system, with a more in-depth discussion of work on category-selective visual responses, particularly face- and body-selective responses. Finally, I will provide an overview and brief summary of each of the five experimental chapters.

Functional magnetic resonance imaging

The main technique used in this thesis is functional magnetic resonance imaging (fMRI). This technique allows for visualisation of brain activity in humans in a safe and non-invasive manner. Although the biological and physiological mechanisms underlying the fMRI signal are not fully understood, there is considerable evidence that the fMRI signal correlates strongly with neural activity (Logothetis et al., 2001; Mukamel et al., 2005; Niessing et al., 2005).

ROI analysis

A common way of analysing fMRI data is to test for each voxel in the brain whether the measured time course of that voxel corresponds to the modelled time course for each of the variables in the experiment. The problem with this approach is that, because of the large number of voxels in the brain, many (>1000) statistical tests are performed simultaneously. As a result, when adopting a significance level of $p < 0.05$, many voxels can be expected to be significantly activated simply by chance (Type 1 errors). Correcting the significance level to reduce Type 1 errors is possible, for instance by using the Bonferroni correction or the false discovery rate (Genovese et al., 2002), but such a correction will increase the number of false negatives, or Type 2 errors. A way around this problem is to specify hypotheses for certain brain regions a priori. By doing so, a more lenient statistical threshold can be used for testing these hypotheses. Another way of reducing the number of statistical tests is by grouping voxels in regions of interest

(ROI). The signals of voxels constituting an ROI are averaged, and only one statistical test for the whole region is performed. ROIs can be defined based on anatomical landmarks (e.g., by combining the response of voxels located in the intraparietal sulcus), or based on functional properties of voxels (e.g., combining all voxels in visual cortex that respond more to moving than to static rings). A further advantage of using individually defined ROIs is that it circumvents variability resulting from individual differences in brain anatomy and functional brain organization that exists even when brains are brought in a common space (Figure 1). The functional ROI approach has been widely used for studying the properties of visual brain areas, as well as the influence of cognitive factors (e.g., attention) on the activity in these areas. However, the functional localization approach relies on the assumption that the location and selectivity of the ROIs are reproducible over time. Stable measurement of ROIs is critical to maximize statistical power and validity in any ROI experiment. Moreover, as some functional brain areas are situated in very crowded neural "neighborhoods", a reliable estimate of the location of ROIs is crucial for the interpretation of experimental results. Chapter 2 investigates this issue, by focussing on the reproducibility of category-selective functional ROIs. These ROIs are used in subsequent chapters. As such, Chapter 2 provides a validation of the use of this method in the other chapters, as well as much of the research that these chapters are based on.

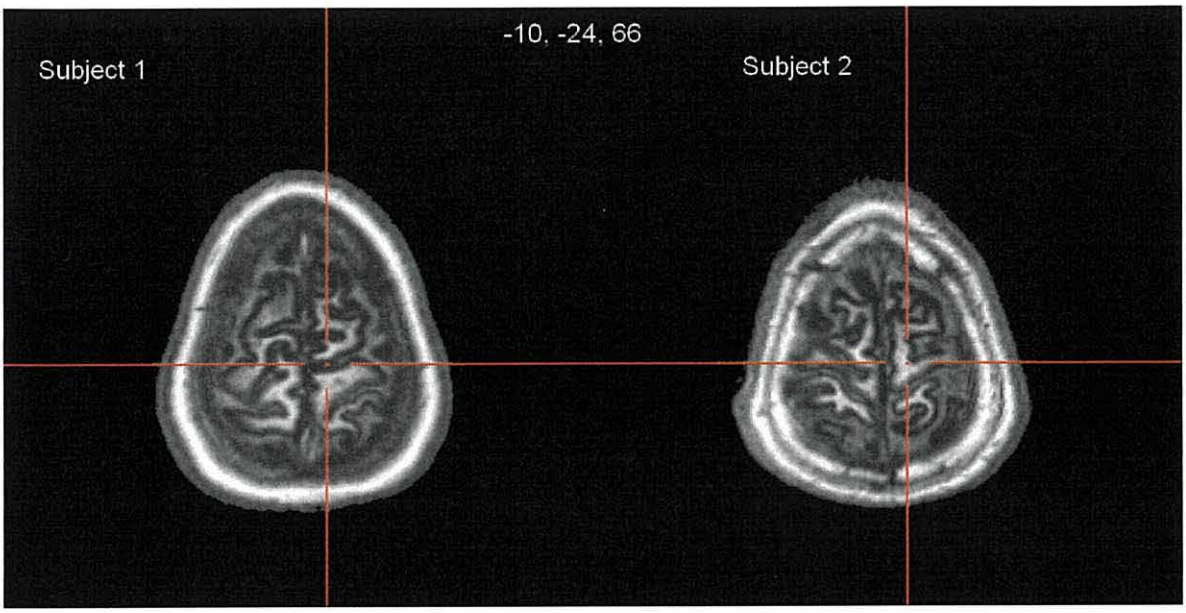


Figure 1. Inter-subject variability in brain anatomy. The same location in Talairach coordinates (-10, -24, 66) displayed on two different, but normalized, brains. In subject 1, the location corresponds to the central sulcus, whereas in subject 2 the location is about one centimetre anterior to the central sulcus.

Analysis of activation patterns

An alternative to testing the strength of activation in a single voxel or ROI is to look at the multivariate pattern of activation across voxels. This approach was first described by Haxby et al., (2001). They showed that the pattern of activation across voxels in the ventral visual stream in response to a visual stimulus was reliable over time. Importantly, this pattern changed as a function of stimulus category, allowing the researchers to predict which object category the subject had viewed based on the pattern of activation at that time. The main focus of the paper was on the extent to which object representations are distributed across higher-level visual cortex. The potential of pattern analysis for addressing other research questions was not fully recognized, however, until

two papers showed that the viewed orientation of simple stimuli can be predicted by the pattern of activation across V1 voxels (Haynes and Rees, 2005; Kamitani and Tong, 2005). This discrimination was not possible by looking at mean activation levels in individual voxels, or groups of voxels. In contrast to Haxby et al., these authors do not interpret their findings as evidence for a distributed representation of orientation in V1. Instead, they argue that each individual V1 voxel contains a mixture of neurons with different orientation tuning. Importantly, the proportion of differently tuned neurons within a voxel is not identical, leading to a slight orientation bias in each voxel. The bias in individual voxels was too small and unreliable to provide information about stimulus orientation. However, when the biases of all V1 voxels were taken into account, using multivariate pattern analyses, stimulus orientation could be reliably predicted. In Chapter 4 we will extend this method and apply it to situations in which functional regions overlap on a relatively fine scale. We show that overlapping regions can be disentangled when using voxelwise pattern analysis.

The visual system

Low-level visual cortex

The human visual system starts with the retina of the eye, on which the visual world is projected through the cornea and lens of the eye. There is a direct (inverted) mapping of the visual world on the retina, resulting in retinal topography (retinotopy). The major output connection of the retina is through the lateral geniculate nucleus (LGN)

of the thalamus to the primary visual cortex (V1) in the occipital lobe. The projections from retina to V1 are crossed, such that the left half of the visual field (from both eyes) is projected onto the right hemisphere, and the right half of the visual field is projected onto the left hemisphere. Area V1 is located in and around the calcarine sulcus. V1 has a retinotopic organization: adjacent locations on the retina activate adjacent locations in V1. The part of V1 on the dorsal bank of the calcarine sulcus contains the lower visual field, whereas the part of V1 on the ventral bank of the calcarine sulcus contains the upper visual field. Thus, in each hemisphere V1 contains a full representation of the contralateral hemifield. V1 neurons have small receptive fields, and relatively simple response properties (e.g., oriented bars). V1 neurons project to neurons in areas V2d (dorsal) and V2v (ventral), the next steps in the visual processing stream. V2 has a retinotopic organization, similar to V1, but receptive fields are slightly larger than V1, and response properties are more complex. This hierarchy continues in the projections from dorsal V2 to area V3 and subsequently area V3a, and ventral V2 to area VP and subsequently area V4v. All these areas are retinotopically organized in a way that makes it possible for fMRI to locate their borders with relative ease (Figure 2; Engel et al., 1997). Beyond V3a and V4v there is less agreement on the functional organization and nomenclature of visual areas.

In sum, the first stage of visual processing takes place in low-level visual areas that are retinotopically organized. There is a hierarchical organization, with neurons in areas further up the processing stream coding for more complex combinations of visual features and having larger receptive fields than areas lower down in the hierarchy.

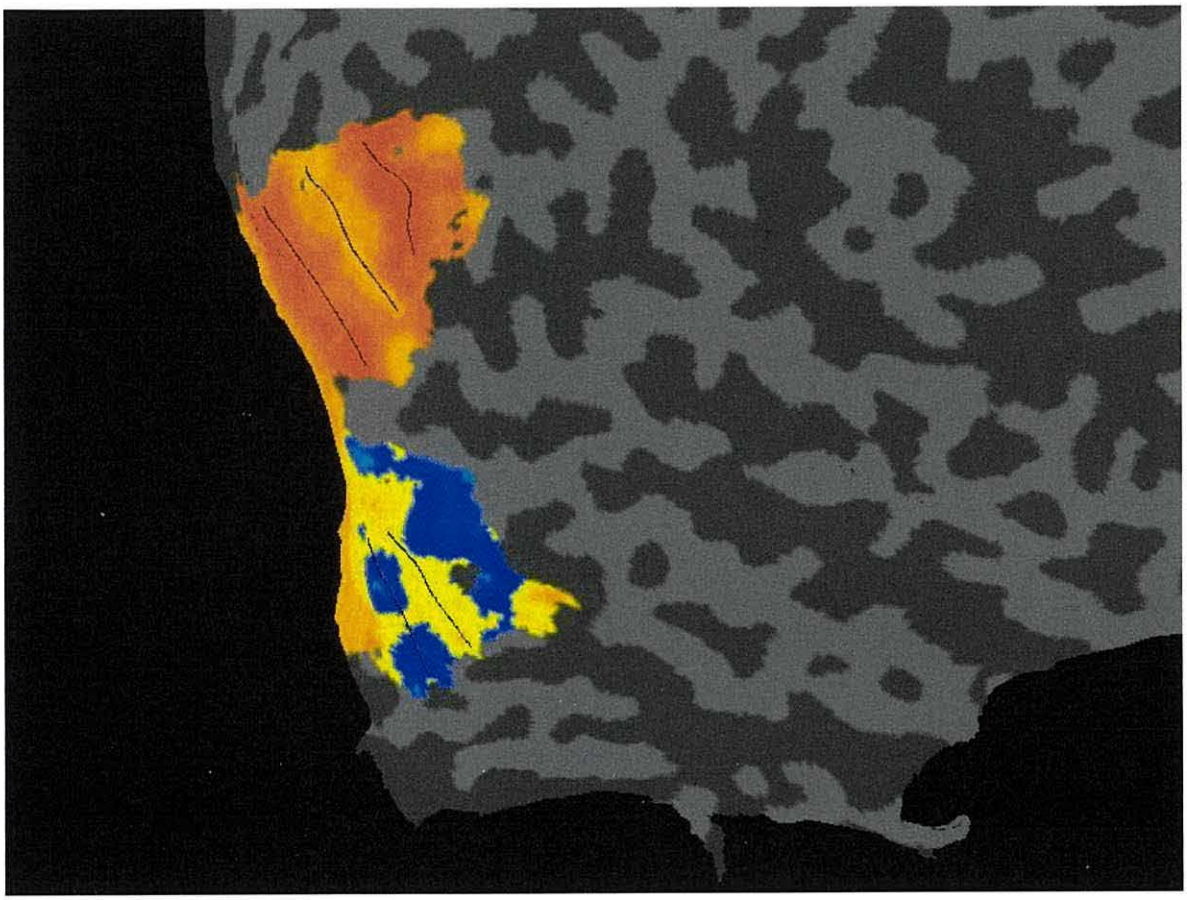


Figure 2. Black lines indicate boundaries between low-level visual areas V1, V2, and V3 of subject PD. The method is based on a visual stimulus that creates a traveling wave of neural activity within retinotopically organized visual areas (Engel et al., 1997).

Higher-level visual cortex

Beyond early retinotopic visual cortex several visual areas with distinct functions have been identified. Among these are motion-sensitive area MT+, and object-selective lateral occipital complex (LOC), which will be discussed next. Visual areas that respond selectively to certain object-categories are discussed separately in another paragraph, given their relevance to the remaining chapters of this thesis.

Area MT+ is located on the posterior end of the inferior temporal sulcus, and is selectively activated by moving versus stationary stimuli (Tootell et al., 1995). There is some evidence, based on retinotopic organization and receptive field size, for a subdivision of the human MT+ complex into two separate areas that may be homologues to macaque areas MT and MST (Huk et al., 2002). In this thesis, however, we will consider area MT+ as a whole, given the complexity of separating the two potential subdivisions.

Area LOC is a large area that is defined by its selectivity to intact versus scrambled object images (Grill-Spector et al., 1998, Malach et al., 1995). It contains two spatially segregated subdivisions: a posterior/dorsal and an anterior/ventral subdivision. The posterior/dorsal part of LOC (sometimes called LO) partially overlaps area MT+ (Kourtzi et al., 2002). The anterior/ventral part overlaps with face-selective areas in the posterior fusiform gyrus. LOC is involved in 2D and 3D shape processing independent of low-level visual features, and is activated by form regardless of the cue (motion, texture or contrast) that defines the form (Grill-Spector et al., 1998; Kourtzi and Kanwisher, 2000; 2001). The anterior/ventral part of LOC is to some extent invariant to the position, size, viewpoint, and orientation of objects (Grill-Spector et al., 1999; Kourtzi et al., 2003; Vuilleumier et al., 2002). As such, LOC can be seen as a homologue of monkey inferior temporal cortex (IT), which contains neurons that code for complex shapes across (small) changes in position, size and orientation (Logothetis et al., 1994; 1995).

Anterior to the anterior/ventral part of LOC is a stretch of ventral cortex bounded by the fusiform gyrus laterally and the parahippocampal gyrus medially that has been referred to as ventral occipital cortex (VOT; Grill-Spector and Malach, 2004). It is in this

part of the visual system that highly face- and place-selective responses have been found (see next paragraphs). Malach and colleagues (Hasson et al., 2002; Levy et al. 2001) proposed that the center-periphery bias seen in early visual cortex extends to VOT: lateral visual areas, including the fusiform gyrus, are thought to be important for center-biased visual field representations (e.g., faces), whereas more medial areas are thought to be important for periphery-biased field representations (e.g., scenes and buildings).

Category selectivity in visual cortex

To what extent is higher-level visual cortex organized along category-specific boundaries? In a recent study, we looked at the response to 20 different object-categories to test whether the cortical structures responsible for human visual recognition contain category-specific regions, each specialized for the perceptual analysis of a distinct class of stimuli (Downing et al., in press). We confirmed previous findings that showed selectivity for faces in the fusiform gyrus (fusiform face area [FFA]; Kanwisher et al., 1997), for places in the parahippocampal gyrus (parahippocampal place area [PPA]; Epstein et al., 1998), and for bodies in the inferior temporal sulcus (extrastriate body area [EBA]; Downing et al., 2001a). Each of these regions responded significantly more strongly to its preferred category than the second most effective category. A region in the middle temporal gyrus that has been reported to respond significantly more strongly to tools than to animals (Chao et al., 1999) did not respond significantly more strongly to tools than to other non-tool categories (such as fruits and vegetables), casting doubt on the characterization of this region as tool selective. We did not find evidence for other

category-selective areas, suggesting that category-selectivity in the visual system is limited to a few categories, at least at the spatial resolution available. In the following sections I will discuss the three aforementioned category-selective regions in more detail.

A challenge to the importance of category-selective areas comes from work showing that distributed patterns of activation outside category-selective areas can reliably predict which category is being viewed (Haxby et al., 2001, but see Spiridon and Kanwisher, 2002). According to Haxby et al., highly selective peaks, such as those seen for faces, places, and bodies, play no unique role in the representation of these objects. Instead, on this view these peaks form part of a broad neural network that represents objects by a distributed collection of feature analyzers. However, this account seems less plausible in light of recent studies that provide evidence that category-selective activation peaks do have a privileged role in object perception. One study correlated fMRI activation with behavioral performance measures (Grill-Spector et al., 2004), and showed a positive correlation between successful face detection and FFA activity. Another study showed impairment in body recognition following TMS over EBA (Urgesi et al., 2004). Finally, a microstimulation study in monkeys showed a direct relation between activation in face-selective cells and face perception (Afraz et al., 2006).

The fusiform face area

Numerous fMRI studies from labs across the world have shown strongly selective responses to faces (including schematic faces) in an area on the fusiform gyrus labelled the FFA (Kanwisher et al., 1997). Control categories included different object types

(Puce et al., 1995, Downing et al., in press), non-face body parts (Kanwisher et al., 1997), views of the back of the head (Tong et al., 2000), and animals (Kanwisher et al., 1999, Chao et al., 1999). The FFA does not respond to faces and other objects in an all-or-none fashion. Instead there is a systematic variation in the response to non-face stimuli, with the strongest (non-face) response to human bodies without heads (Downing et al., in press). Downing and colleagues suggested that this may reflect "partial voluming" effects: pooling across distinct neural populations that are interleaved within a voxel but that each respond exclusively to a single stimulus category. An imaging studying using high spatial resolution (1.4 X 1.4 X 2.0 mm) has confirmed this explanation, at least for the case of bodies in the FFA, by showing very small, distinct patches of cortex that were either selective for bodies but not faces, or for faces but not bodies (Schwarzlose et al., 2005). Chapter 3 and 4 will provide further evidence supporting this explanation.

Face selective cells have also been found in monkeys using single-cell recordings (Desimone et al., 1984; Gross et al., 1972; Perrett et al., 1992). These cells are mostly located on the ventral bank of the superior temporal sulcus and on the polysensory area on the dorsal bank of STS. Recent studies using fMRI suggest that these cells are clustered together, and that they may be homologues to human FFA (Pinsk et al., 2005, Tsao et al., 2003).

Relation between FFA and perception

What is the function of the FFA? Does its activity correlate with the subjective percept of a face? Does it subserve face recognition, face identification, or both? Numerous studies have found a relation between activity in the FFA and the subjective percept of a face, e.g., in the vase-face illusion (Andrews et al., 2002), the perception of Mooney faces (Andrews and Schluppeck, 2004), or during binocular rivalry (Tong et al., 1998). The FFA has also been implicated in face imagery (O'Craven and Kanwisher 2000), and activity in the FFA can be modulated by attention (Avidan et al., 2003; O'Craven et al., 1999; Wojciulik et al., 1998). Together these studies provide strong evidence for a link between FFA activity and the subjective awareness of seeing a face. Evidence for a role of the FFA in face detection and identification comes from a study measuring the correlation between FFA activity and psychophysical performance on detection and identification tasks (Grill-Spector et al., 2004). The results showed that successful detection of a face (but not successful detection of objects) correlated with FFA activity. Even stronger activity was seen when faces were also correctly identified. This pattern of results was not seen in other visual areas, suggesting that the FFA is indeed critically involved in both detection and identification of faces. In monkeys, stimulation of face-selective cells has been shown to influence face categorization, such that they were more likely to report seeing a face in a noise display when microstimulation was applied to face-selective cells compared to when microstimulation was applied to non-selective cells (Afraz et al., 2006). These results suggest a direct link between activation in face cells and face perception.

Are faces special?

Why is the FFA selective for faces? One account is that faces constitute a special object class, and are processed differently than other object classes (Yin, 1969). On this account the response in the FFA to non-faces reflects a partial (but epiphenomenal) engagement of a truly face-selective module (Kanwisher, 2000).

An alternative account argues that the selectivity for faces in FFA can be entirely explained by the experience we have with discriminating faces (Tarr and Gauthier, 2000). According to this account, face-selective neurons in FFA are involved in a specific cognitive process, independent of stimulus category: distinguishing among structurally similar exemplars of a category for which one has substantial expertise. Evidence in favour of this account includes the finding that car and bird experts show enhanced FFA activity when viewing objects from their category of expertise (Gauthier et al., 2000). Another study trained subjects to learn to recognize "greebles", an artifical object category specifically devised for testing the expertise account. Subjects showed enhanced FFA activity after extensive training in recognizing greebles (Gauthier et al., 1999). It must be noted, however, that the effects of expertise in these studies is small, and faces still give a much stronger response than objects of expertise. Secondly, greebles might not be ideal for testing the expertise account, as they have the same basic configuration as faces (symmetrical configuration in which two horizontally arranged parts are above two vertically aligned central parts). Finally, in these studies subjects are trained to recognize greebles using proper names, which might induce an animate or human interpretation of

these stimuli (Kanwisher, 2000). Future studies should test for the effects of expertise with stimuli that are not face-like and that are encoded as inanimate objects (e.g., tools), or could systematically vary these variables to test the influence of these factors on activity in the FFA.

It is likely that neurons in higher-level visual cortex have some degree of plasticity, such that neurons can become tuned to respond holistically to objects of expertise (e.g., in order to process these stimuli more efficiently or to discriminate exemplars more accurately). The studies investigating the effects of expertise suggest that these neurons are located in the fusiform gyrus, including FFA. It seems improbable, however, that the same neuron would code for both faces and objects of expertise (e.g., cars), as argued by Gauthier and colleagues. Indeed, a recent monkey study combined fMRI and single-cell recordings to show a cluster of neurons, in a region that may correspond to FFA in humans, with exclusively selectivity for faces (Tsao et al., 2006). Thus, faces may be special in that many neurons, clustered together, are tuned to respond selectively to faces (enough to show up in fMRI). In Chapters 3 and 4 we describe the finding that bodies (without heads) also strongly activate the fusiform gyrus in a region that overlaps with but is distinct from FFA. We interpret this as reflecting overlapping populations of highly selective neurons that co-exist at close quarters. Faces and bodies are represented by larger populations and/or more selective neurons than other categories, which may well be due to the extensive experience we have with these categories.

The parahippocampal place area

Another proposed category-selective area in visual cortex is the parahippocampal place area (PPA; Epstein and Kanwisher., 1998). This area, located in the parahippocampal cortex, responds selectively to images of indoor and outdoor scenes, independently of the complexity of those scenes (it responds equally strongly to two walls delineating a room as it does to completely furnished rooms). Its selectivity for scenes is striking, with very weak responses to other stimulus categories (Downing et al., in press). A possible role of this area is the encoding of scenes into memory, possibly for later orientation (Epstein et al., 2001). Another study implicated the PPA in egocentric rather than allocentric representations of space (Epstein et al., 2003). The precise role of the PPA in navigation, and its interactions with other structures such as the hippocampus is at present unclear.

The extrastriate body area

The extrastriate body area (EBA; Downing et al., 2001a) is an area located on the posterior end of the inferior temporal sulcus that responds generally and selectively to images of bodies (without heads) as compared to a wide variety of control stimuli, including faces (Downing et al., 2001a; in press). Its response is unlikely to be related to low-level visual features of body stimuli, as the EBA also responds selectively to schematic bodies (e.g., stick figures or silhouettes, versus visually similar controls). Anatomically, the EBA overlaps with, but can be dissociated from, motion-selective

MT+ and object-selective LO (Downing et al., 2001a). The EBA does not overlap with the FFA (Figure 3). Finally, a TMS study showed selective impairment of body recognition after disruption of the EBA (Urgesi et al., 2004), suggesting that the EBA is necessary for successful recognition of bodies. In monkeys, hand-selective cells have been reported (Wachsmuth et al., 1994), and monkey fMRI studies have shown a body-selective area on the ventral bank of the STS, neighbouring a face-selective area (Pinsk et al., 2005). Whether these areas correspond to human EBA is unclear. Below I will review some of the current debates surrounding the possible functions of the EBA.

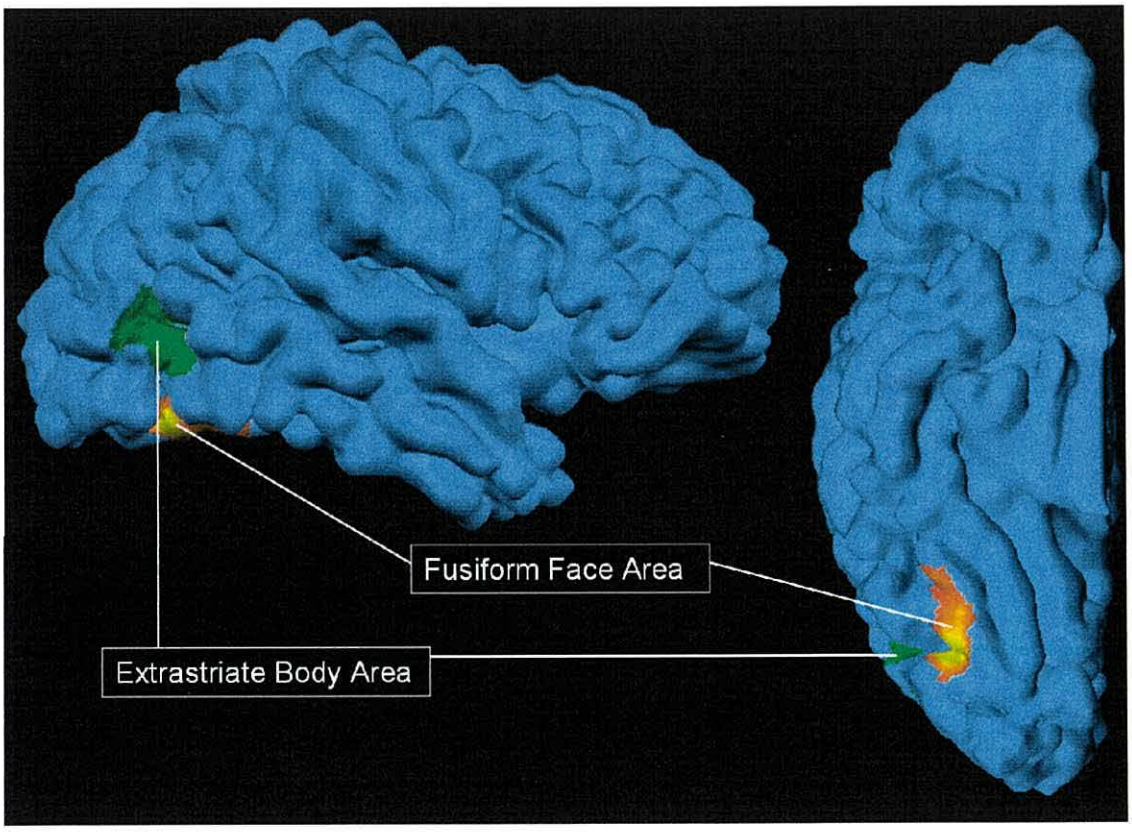


Figure 3. Location of EBA and FFA in one subject. Note that EBA and FFA do not overlap: EBA is located on the posterior end of the inferior temporal sulcus, whereas FFA is located more ventrally in the region of the fusiform gyrus.

Role of the EBA in processing biological motion

Given its close proximity to MT+, a primary function of the EBA could be the processing of moving bodies and body parts. Movement of the body might even be necessary to activate EBA. On this account, the selectivity for static bodies could be explained by implied motion of images of bodies, similar to the finding of selective activation in MT+ to images that imply simple motion (Kourtzi and Kanwisher, 2000). To test this, Downing et al. (2001a) compared the response in the EBA to static images of active, awake humans (that could imply motion) with the response to static images of sleeping humans. Equally strong body-selectivity was found for both conditions, suggesting that real or implied motion is not necessary for activation of the EBA. However, several studies have reported significant activation in the EBA to point-light displays of biological motion (Johansson, 1973) versus scrambled controls (Downing et al., 2001a; Grossman and Blake, 2002; Michels et al., 2005). Does this finding provide evidence for a role of the EBA in processing body movements? This question will be addressed in Chapter 4 and 5. In Chapter 4 we will investigate whether the previously observed EBA activation to biological motion displays can be explained by activation of overlapping area MT+. In Chapter 5 we look at selectivity for body movements in a more controlled way than previous studies. Specifically, we measure the EBA's response to coherent, meaningful body movements versus incoherent, meaningless body movements. This provides a measure of movement selectivity without confounding the perception of a human body, as was the case in studies using point-light displays.

Role of the EBA in the execution of visually guided action

A recent study (Astafiev et al., 2004) found that the EBA is not only involved in the visual recognition of bodies but also in the execution of visually guided (unseen) movements of body parts (hands and feet). The authors suggest that such a movement may affect the actor's body representation through proprioceptive inputs that result from the movement. Alternatively, or additionally, the EBA could be activated through corollary discharge signals originating in motor areas. These signals may serve to dynamically update body representation in EBA, and adjust for sensory input resulting from the movement. Ultimately, these signals could function to distinguish between one's own and someone else's body (Jeannerod, 2004; see also next paragraph). In Chapter 6 we replicated this experiment and used several analyses to test this notion. These more careful analyses contradicted the proposal of Astafiev et al. (2004) that the EBA is involved in motor action.

Self or other?

Is the EBA primarily involved in perceiving other people, or is its main function the perception of one's own body? Two studies tested the response in the EBA to allocentric and egocentric pictures of bodies (Chan et al., 2004) and body parts (Saxe et al., 2006). In both studies right EBA responded more to allocentric compared to egocentric views. No difference was found when comparing the response to images of the observer's own body or another familiar person (Chan et al., 2004). Together, these

results argue against a primary role for the EBA in perceiving the self or in guiding one's actions. Instead, the bias for allocentric views of bodies in right EBA could signify a first stage of a social vision network responsible for perceiving other people.

Overview of thesis

Chapter 2: Within-subject reproducibility of category-selective visual activation with functional MRI

This chapter describes the results of a study investigating the reproducibility of fMRI activation in higher-level visual cortex. It focuses on category-selective regions in ventro-temporal cortex, and provides a validation of the methods used in the other chapters.

In a first scanning session, subjects (N=6) were presented with visual images of bodies, faces, scenes, and tools. By contrasting fMRI activation in response to these images, we identified several previously described category-selective regions: the body-selective extrastriate body area (EBA), face-selective fusiform face area (FFA), and scene-selective parahippocampal place area (PPA). In a second scanning session, three weeks after the first session, the experiment was repeated in the same subject group. On an individual-subject basis, we compared the activation in session one to that in session two and found a striking similarity between the results of the two sessions. On average, the difference between the two sessions in the location of activation peaks was 2.9 mm, which is below the typical scanning resolution. Functional reproducibility, defined by the stability of selectivity in the areas, was also high.

We conclude that category-selective regions such as the EBA, FFA, and PPA can be robustly localized in individual subjects. This provides a validation of the methods used in the remainder of this thesis, as well as much of the research that it is based on.

Chapter 3: Selectivity for the human body in the fusiform gyrus

This chapter describes the discovery of an area in the fusiform gyrus that responds generally and selectively to images of the human body.

Twenty two subjects were scanned while they watched images of bodies, faces, scenes, and tools. As expected, contrasting bodies with tools resulted in activation in the EBA. Unexpected was a second focus of activation in the fusiform gyrus. We compared the location of this area with the face-selective FFA, and found that the two areas overlapped considerably. Detailed analyses of the distance between the two areas and their functional profile, however, showed that the two areas were not identical. In a second experiment (N=8), we tested whether the body-selectivity of this fusiform body area (FBA) extended to abstract images of the human form (stick-figure representations). The FBA (and EBA), but not the FFA showed significant selectivity for these figures as compared to matched control images.

The finding of body-selectivity in the fusiform gyrus challenges accounts of this region that focus solely on faces, and suggests that this region contains multiple distinct category-selective neural representations.

Chapter 4: Patterns of fMRI activity dissociate overlapping functional brain areas that respond to biological motion

This chapter uses voxelwise pattern analysis to show that responses to biological motion in visual cortex can be attributed to body-selective regions instead of overlapping face- and motion-selective regions.

We localized several visual areas (FBA, FFA, EBA, MT+) in 18 subjects, and measured the response in these areas to displays of "point-light" biological motion. FFA and FBA overlapped substantially, replicating the results of Chapter 3. Similarly, EBA and motion-selective MT+ showed considerable overlap. Due to this overlap, all areas showed a significant response to biological motion compared to control displays, even when just the peak voxels of these areas were considered. Strikingly, however, only body selectivity was correlated, on a voxel-by-voxel basis, with biological motion selectivity.

We conclude that biological motion, through the process of structure-from-motion, engages areas involved in the analysis of the static human form. Furthermore, these results show that body-selective regions in posterior fusiform gyrus and posterior inferior temporal sulcus overlap with, but are functionally distinct from, face- and motion-selective regions.

Chapter 5: The role of the extrastriate body area in action perception

In Chapter 4 we showed that the EBA, FBA, and superior temporal sulcus respond significantly to displays of biological motion. Are these areas all processing biological movements, or is the response to point-light displays in EBA and FBA due to structure-from-motion information that is present only in the intact condition?

To address this question we presented two kinds of stimulus sequences. In the *coherent* condition, static frames from a movie of a single, intransitive whole-body action

were presented in the correct order. In the *incoherent* condition, a series of frames from multiple actions (involving one actor) were presented. By presenting the same frames but in different order, there was no confound in form information between the two conditions. Thus, this paradigm allowed us to test for regions that are involved in constructing a dynamic representation of the observed action. ROI analyses (N=16) showed that the EBA, unlike area MT+ and the posterior superior temporal sulcus, responded more to the incoherent than to the coherent conditions.

We interpret this finding as evidence for a unique role of the EBA in the perception of action, by representing the static structure, rather than dynamic aspects, of the human form.

Chapter 6: The role of the extrastriate body area in motor actions

Astafiev et al. (2004) reported that the EBA is not only involved in the visual recognition of bodies but also in the execution of visually guided (unseen) movement of body parts (hands and feet). Here we replicated their experiment.

We found, in line with Astafiev et al., action-related modulation in the EBA. However, a closer look showed that the region involved in visually guided motor acts only marginally overlapped with EBA. Furthermore, action-related modulation and body-selectivity were unrelated on a voxel-by-voxel basis. That is, voxels that showed strong body-selectivity did not necessarily show strong action related modulation. We conclude that the temporo-occipital area that is involved in executing motor actions is distinct from the EBA.

Chapter 2:

Within-subject reproducibility of category-specific visual activation with functional MRI

ABSTRACT

The present study used fMRI to investigate the within-subject reproducibility of activation in higher level, category-specific visual areas in order to validate the functional localization approach widely used for these areas. The brain areas we investigated included the extrastriate body area (EBA), which responds selectively to human bodies, the fusiform face area (FFA) and the occipital face area (OFA), which respond selectively to faces, and the parahippocampal place area (PPA), which responds selectively to places and scenes. All 6 subjects showed significant bilateral activation in the four areas. Reproducibility was very high for all areas both within a scanning session and between scanning sessions separated by 3 weeks. Within sessions, the mean distance between peak voxels of the same area localized by using different functional runs was 1.5 mm. The mean distance between peak voxels of areas localized in different sessions was 2.9 mm. Functional reproducibility, as expressed by the stability of T-values across sessions, was high for both within-session and between-session comparisons. We conclude that, within subjects, high-level category-specific visual areas can be localized robustly across scanning sessions.

INTRODUCTION

FMRI research on visual perception has revealed several distinct bilateral occipito-temporal brain areas that respond selectively to certain categories of visual stimuli. These include: the extrastriate body area (EBA), which responds selectively to human bodies and body parts (Downing et al., 2001a), the fusiform face area (FFA) and the occipital face area (OFA), which respond selectively to faces (Halgren et al., 1999; Kanwisher et al., 1997; Puce et al., 1996), and the parahippocampal place area (PPA), which responds selectively to places and scenes (Epstein and Kanwisher, 1998).

The discovery of these category-selective visual areas has initiated further research into their response properties (e.g., Chan et al., 2004; Chao et al., 1999; Epstein et al., 2003; Kanwisher et al., 1999; Tong et al., 2000), their role in cognitive functions like working memory (Druzgal and D'Esposito, 2003) and imagery (Ishai et al., 2002; O'Craven and Kanwisher, 2000), and factors that may modulate their activity, such as attention (Avidan et al., 2003; O'Craven et al., 1999; Wojciulik et al., 1998) and familiarity (Epstein et al., 1999; Rossion et al., 2003). These areas have also proven useful for testing cognitive models of visual attention (Downing et al., 2001b; de Fockert et al., 2001; O'Craven et al., 1999, Yantis and Serences, 2003).

However, it must be noted that category-specific responses in a region do not necessarily indicate a special role for that region in object recognition. Indeed, according to one account of ventral stream organization, these selective "peaks" simply form one part of a large distributed network, and do not contribute more or less than other regions to object recognition (Haxby et al., 2001; but see Spiridon and Kanwisher 2002).

Recently, however, it has been shown that activity in the FFA correlates with successful detection and identification of faces, but not with other stimulus categories (Grill-Spector et al., 2004). This correlation was not observed in other ventral visual areas, even those (such as the PPA) where the response to faces is consistently low and could in principle provide information relevant to face processing. Thus, although this debate remains to be conclusively resolved, there is at least some evidence that category-selective peaks have a privileged role in vision.

A widely-used strategy for studying the properties of category-specific visual areas, or the influence of cognitive factors on these areas, is to functionally define the region of interest (ROI) for each subject with a localization experiment (e.g., by contrasting faces with other objects to define the FFA). Within these ROIs, effects of subsequent experimental manipulations can then be measured. The localizer measurement and the experimental manipulation of interest are usually performed in separate runs of the same fMRI session. In some cases it would be advantageous to perform the localizer in a different scanning session, for three reasons: (1) more time is left for the experiment of interest, allowing the inclusion of more conditions and/or increased power; (2) when testing the same group of subjects on multiple experiments involving the same ROIs, localization can be performed only once for each subject; and (3) localizing across sessions makes it possible to test the response in newly localized areas to experimental conditions of previously acquired data.

The functional localization approach relies on the assumption that the location and selectivity of the ROIs are reliable across runs, either within or between scanning sessions. Stable measurement of ROIs is critical to maximize statistical power and

validity in an ROI experiment. Moreover, as some of the category-specific areas mentioned above are situated in very crowded neural "neighborhoods" (e.g., the EBA partially overlaps MT and the dorso-lateral focus of LOC [Downing et al., 2001a; Malach et al., 1995]) a reliable estimate of the location of ROIs is crucial for the interpretation of experimental results. Thus, the aim of the present study was to establish an estimate of the within-subject reliability of category-specific visual areas, within and between scanning sessions.

Previous work investigating within-subject reproducibility of fMRI activation has focused on activation in primary visual cortex (Miki et al., 2000; Miki et al., 2001a, 2001b; Rombouts et al., 1998), motor cortex (Ramsey et al., 1996; Yetkin et al., 1996; Tegeler et al., 1999), somatosensory cortex (Yetkin et al., 1996) and medial temporal lobe (Machielsen et al., 2000). These studies assessed the reproducibility of *all* of the activation produced by some stimulus or task, rather than the reliability of specific, focal *a priori* regions of interest. The measure of reproducibility used in these previous studies tested the amount of spatial overlap of activation across different data sets. Because we were interested in the reliability of the location and selectivity of category-specific ROIs, we calculated the following measures of reproducibility: (1) the distance between peak voxels, and (2) the reproducibility of statistical values within the ROI.

MATERIALS AND METHODS

Participants

Six healthy volunteers (5 male, age range 21-33 years) with normal or corrected-to-normal vision participated in the experiment. All subjects gave informed consent, and experimental procedures were approved by the ethical board of the School of Psychology at the University of Wales, Bangor and the North-West Wales Health Trust.

Experimental paradigm

Subjects were scanned on two occasions (session A and B) separated by 21 to 23 days (median 21 days). The stimulation protocol was identical in the two sessions. No technical or software updates were performed between the two scanning sessions. Each session consisted of 4 runs. Each run consisted of 21 15-second blocks. Five of these 21 blocks were fixation-only baseline conditions, occurring on block 1, 6, 11, 16 and 21. During the other 16 blocks, subjects were presented with pictures of faces, bodies, tools or scenes. Forty full color exemplars of each category were tested (see Figure 1 for examples). These were divided into two sets of twenty stimuli, which were presented during runs one and two, and runs three and four respectively. The order of blocks was symmetrically counterbalanced within each run. Two versions of the block order were adopted. The first half and second half of one version were swapped to create the second version. The block order of the first run matched the block order of the third run, and the same was true for the second and fourth runs. In the picture blocks, each stimulus was presented for 300 msec, with an ISI of 450 msec. Stimuli were back-projected on the

center of a screen, which was viewed by the subjects through an angled mirror positioned on top of the head coil. Subjects had to press a button whenever a stimulus appeared twice in immediate succession (a 'one back' task), which happened twice per block.

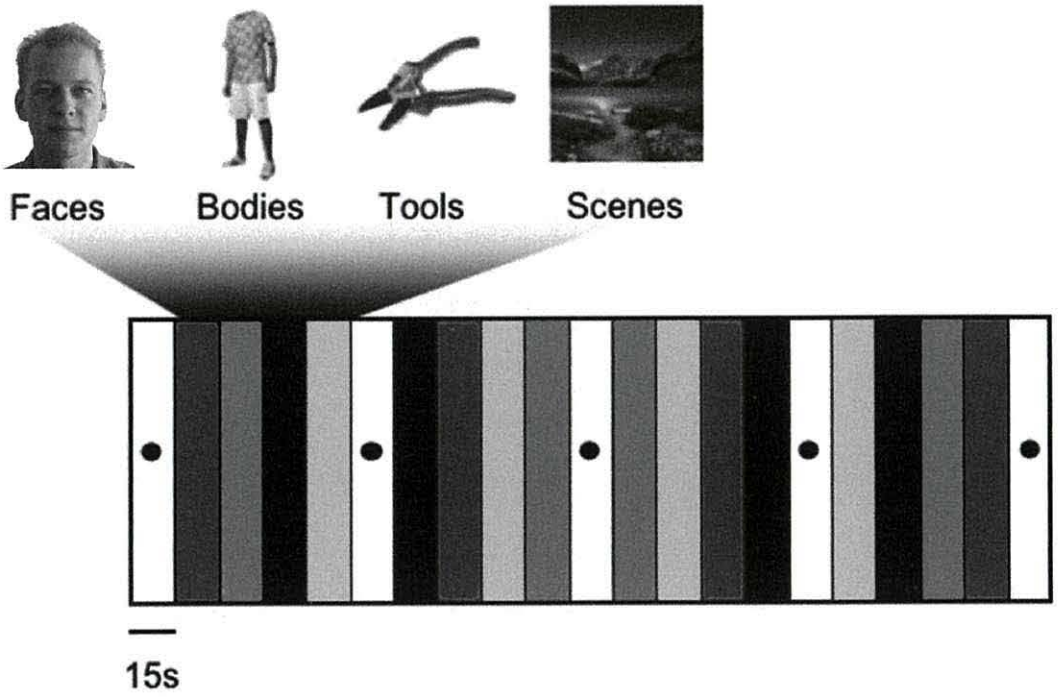


Figure 1. Example stimuli and overview of the design.

Functional imaging and analysis

Brain imaging was performed on a Philips Gyroscan Intera 1.5 T scanner equipped with a SENSE head coil (Pruessmann et al., 1999). An EPI sequence was used to image functional activation. Within subjects, the same scanning protocol was used in both sessions. Thirty oblique slices were collected per image covering the whole brain. Scanning parameters were: repetition time / echo time (TR/TE) = 3000/50 ms, flip angle (FA) = 90°, slice thickness = 4 or 5 mm (no gap), acquisition matrix = 64 X 64, in-plane

resolution = 4 X 4 mm. For anatomical localization, a structural scan was made for each subject using a T1-weighted sequence. Scanning parameters were: TR/TE = 11.5/ 2.95 ms, FA = 8°, coronal slice thickness = 1.3 mm, no gap, acquisition matrix = 256 X 256, in-plane resolution = 1 X 1 mm.

Pre-processing and statistical analysis of fMRI data were performed using BrainVoyager 4.9 (Brain Innovation, Maastricht, The Netherlands). The first three volumes of each run were discarded in order to avoid differences in T1 saturation. The first volume of each functional run was aligned to the first volume of the first functional run (intra-session alignment). Functional data were motion-corrected, low-frequency drifts were removed with a temporal high-pass filter (0.006 Hz), and the data were spatially smoothed with a Gaussian kernel (FWHM 6 mm). Functional data were then manually co-registered with the individual 3-D anatomical scans. The 3-D scans were transformed into Talairach space (Talairach and Tournoux, 1988), and the parameters for this transformation were subsequently applied to the co-registered functional data. Voxel time courses were re-sampled to a resolution of 3 X 3 X 3 mm using trilinear interpolation. Note that spatial normalization is unlikely to affect within-subject reliability (Swallow et al., 2003).

In order to generate predictors for the multiple-regression analyses, the event time series were convolved with a delayed γ function (δ = 2.5 s; τ = 1.25 s) in order to model the hemodynamic response (Boynton et al., 1996). Voxel time series were *z*-normalized for each run, and additional predictors accounting for baseline differences between runs were included in the design matrix.

ROIs (EBA, FFA, OFA, and PPA) were defined for each subject by performing 3 multiple-regression analyses. To localize the EBA, activation caused by images of bodies was contrasted with the average activation of the remaining 3 stimulus categories. Similarly, faces were contrasted with the other categories to localize the FFA and OFA, and scenes were contrasted with the other categories to localize the PPA. For each contrast we located the most significantly activated voxel near previously described locations of the ROIs (EBA: posterior end of the inferior temporal sulcus [Downing et al., 2001a]; FFA: mid-fusiform gyrus [Kanwisher et al., 1997]; OFA: inferior occipital sulcus [Puce et al., 1996]; PPA: parahippocampal gyrus [Epstein and Kanwisher, 1998]). ROIs were defined as the set of contiguous voxels that were significantly ($p<0.0001$, uncorrected for multiple comparisons) activated within 8 mm in the anterior/posterior, superior/inferior, and medial/lateral direction of the most significantly activated voxel. We limited the number of included voxels instead of including all contiguous voxels that are activated at a certain threshold because the group of contiguous voxels that are activated by a contrast can be very large and can extend up to the primary visual cortex in the case of scenes (due to the larger image size of scenes compared to the other categories). Additionally, for the contrast involving faces, the FFA and OFA often merge together, only being separated at a certain threshold that differs for each subject and for the two hemispheres within a subject. Therefore, to avoid setting different thresholds for each subject and ROI, we selected only the group of significantly-active voxels that were within close proximity to the peak voxel, ensuring that only the most selective voxels were included.

Reproducibility measures

We assessed both the within-session and between-session reproducibility of the activation in the EBA, FFA, OFA, and PPA. For between-session reproducibility, we compared the activation in session A with the activation in session B. To create a comparable within-session comparison, we combined run 1 and run 3 of session A with run 2 and run 4 of session B. This dataset will be referred to as session AB. Similarly, we combined run 2 and run 4 of session A with run 1 and run 3 of session B. This dataset will be referred to as session BA. Within-session reproducibility was assessed by comparing session AB with BA. In this way, both between- and within-session comparisons had an equal amount of runs, identical stimuli, and the same stimulus orders on both sides of the comparison.

Distance between peak voxels

The first measure of reproducibility concerned the spatial reliability of the ROIs. We calculated the linear distance between the most significantly activated voxels in the different sessions, for each ROI and hemisphere separately. Thus, for the between-session comparison we calculated the distance between peak voxels in session A and B. Similarly, for the within-session comparison we calculated the distance between peak voxels in session AB and BA. A session*ROI*hemisphere repeated-measures ANOVA was performed to test differences in linear distance.

Functional reproducibility

For each session (i.e. A, B, AB, and BA), we defined the category-specific ROIs and T-values of the ensemble response of the voxels in these ROIs. T-values were based on the same statistical contrasts used to define the ROIs. To assess functional reproducibility, we compared the T-values of different runs in the following way. First, the T-value of the ROIs was computed for the runs that were also used for defining these ROIs (e.g., the T-value was computed for the data from session A based on the ROI defined by session A). Then, T-values were computed for the remaining session (e.g., the T-value of session B in the ROI defined by session A). The comparison of the T-values of ROIs that were defined in the same runs with the T-values of ROIs that were defined in different runs gives an indication of the functional reproducibility of the ROIs. If category-selective activations are highly stable, we expect similar T values in both cases. We compared session A and B (between-session comparison) and session AB and BA (within-session comparison). Sign tests were performed to test differences in reproducibility between ROIs.

RESULTS

Category-specific activation

Figure 2 shows the category-specific activations in a representative subject for the 4 different sessions (A, B, AB, BA). As can be seen, bodies (top panel), faces (middle panel) and scenes (bottom panel) selectively activated the hypothesized regions (EBA, FFA, OFA, and PPA). On visual inspection the activations appeared to be very similar across sessions. All areas could be localized bilaterally in all sessions for all 6 subjects. Mean Talairach coordinates (averaged across sessions and subjects) for the regions were: EBA (left: -43, -72, -2; right: 46, -70, -1), FFA (left: -38, -46, -16; right: 41, -47, -17), OFA (left: -36, -73, -17; right: 37, -74, -17), PPA (left: -23, -44, -9; right: 27, -40, -7)

Mean T-values

Figure 3 gives the T-values (averaged across subjects) of the four ROIs for each session and hemisphere. The highest T-values were observed in the PPA and high, but somewhat lower T-values in the EBA, FFA, and OFA.

The mean T-values (averaged across ROIs) of session A (T = 12.3) and B (T = 12.2), and session AB (T = 11.9) and BA (T = 12.4) were comparable. Within subjects, the absolute difference of T-values between session A and B was larger than the absolute difference between session AB and BA (2.1 versus 1.5). Thus, within subjects, T-values differed more between sessions than within sessions, but not in a systematic direction (that is, for some subjects $T_A > T_B$ while for others $T_B > T_A$).

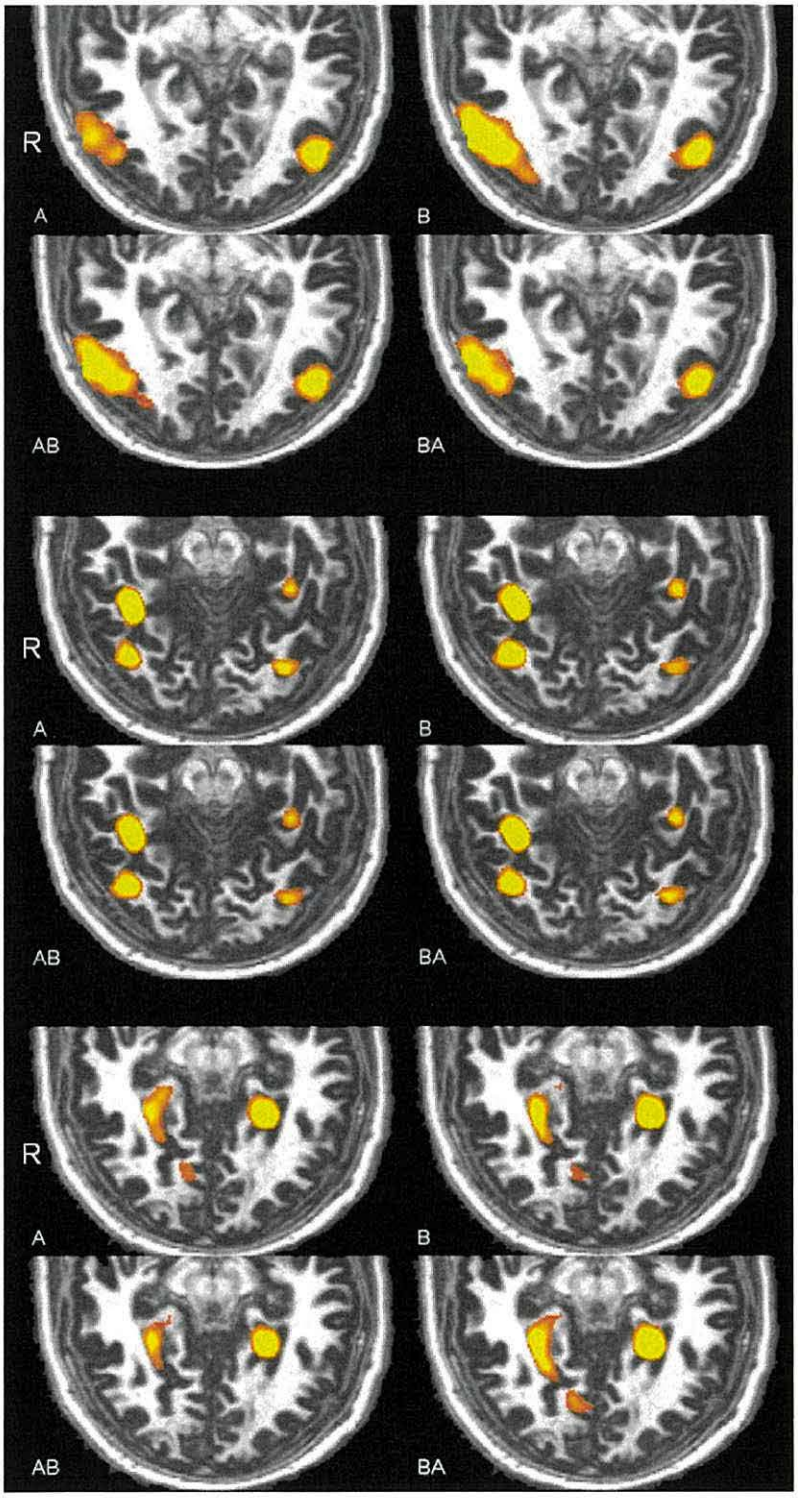


Figure 2. Transverse slices showing category-specific activity for bodies (top panel; z = -4, T > 6), faces (middle panel; z = -18, T > 6), and scenes (bottom panel; z = -8, T > 12) in the four sessions. The left side of the image corresponds to the right side of the brain.

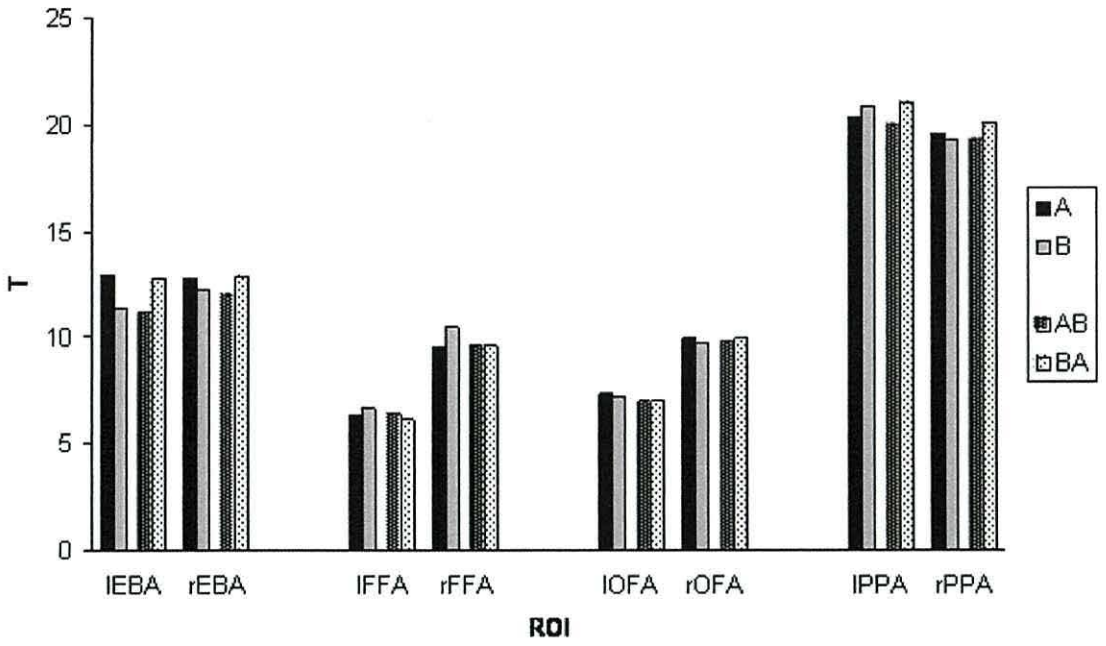


Figure 3. Mean T-values (averaged across subjects) for each ROI in the four sessions. T-values were computed by using the same contrasts as those used for defining the ROIs (e.g., bodies vs. other categories for the EBA).

Distance between peak voxels

Figure 4 shows the distances between peak voxels for each session (within or between), ROI and hemisphere. A 2*4*2 repeated-measures ANOVA (session*ROI*hemisphere) on the distance between peak voxels revealed only a significant effect of session ($F_{1,5} = 12.5$, $p < 0.05$); there was a larger distance between scanning sessions (mean = 2.9 mm, SE = 0.40) than within scanning sessions (mean = 1.5 mm, SE = 0.34). This effect did not depend on the ROI or hemisphere, or the interaction between ROI and hemisphere. That is, there were no significant interactions between

session and ROI (p = 0.18), between session and hemisphere (p = 0.77), or between session and ROI*hemisphere (p = 0.36).

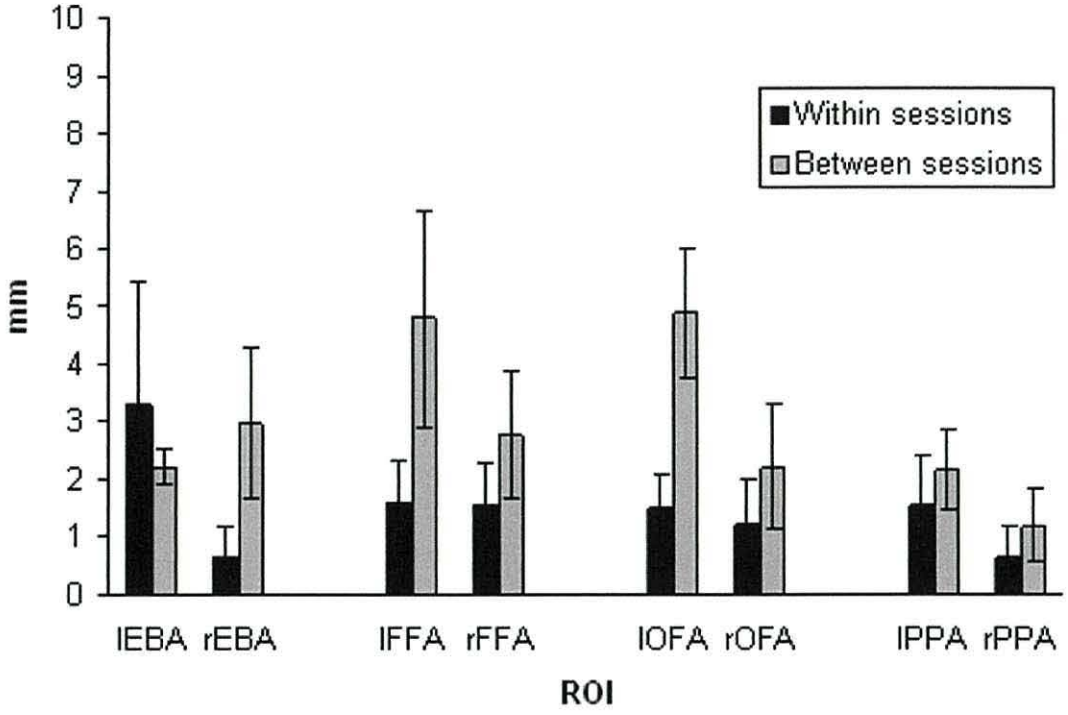


Figure 4. Linear distance between peak voxels for each ROI, for within- and between-session comparisons. Error bars indicate standard errors of the mean.

Functional reproducibility

Figure 5 shows the T-values of the ROIs as a function of the runs in which the ROIs were defined (same or different) for both within- and between-session comparisons. Within sessions, the average T-value when the ROIs were defined in the same runs was 12.2, compared to 12.0 when the ROIs were defined across runs (a reduction of 1.6%). Between sessions, the average T-value when the ROIs were defined in the same runs was

12.3, compared to 11.6 when the ROIs were defined across runs separated by three weeks (a reduction of 5.7%). A sign test comparing the reduction in T between sessions with that within sessions was significant ($p < 0.05$), indicating higher reproducibility within sessions.

The four ROIs (averaged across hemispheres) did not differ in reproducibility (as expressed by the reduction in T-value) for the within-session comparisons (all $ps > 0.10$). Between sessions, we found higher reproducibility for the PPA compared to both the FFA and the OFA (both $ps < 0.05$).

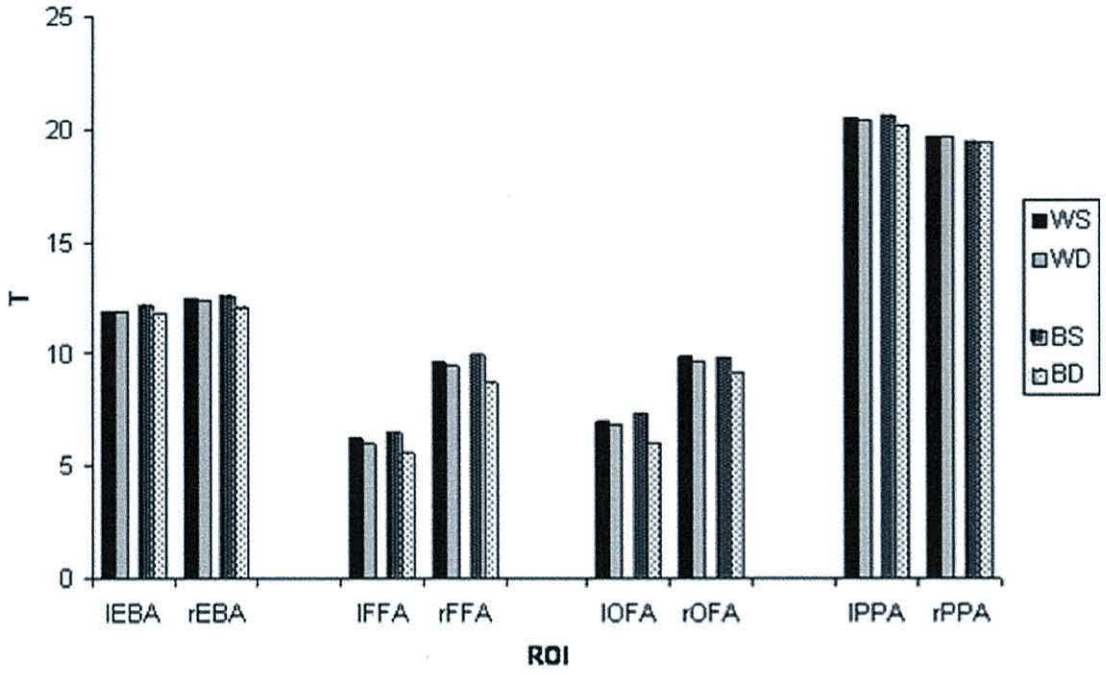


Figure 5. Mean T-values (averaged across subjects) for each ROI as a function of whether the ROI was defined in the same or different runs, for within- and between-session comparisons. WS = within-session, same runs; WD = within-session, different runs; BS = between-session, same runs; BD = between-session, different runs.

DISCUSSION

The present study investigated the within-subject reproducibility of category-specific ROIs in ventral occipito-temporal cortex. As expected, pictures of bodies selectively activated the EBA (Downing et al., 2001a), faces activated the FFA and OFA (Halgren et al., 1999; Kanwisher et al., 1997; Puce et al., 1996), and scenes activated the PPA (Epstein and Kanwisher, 1998). All 6 subjects showed significant bilateral activation in all ROIs.

To test one of the key assumptions of the functional localization method -- that the location and selectivity of ROIs are stable -- we assessed the reliability of ROI definition by looking at the consistency of activation across runs. Within a scanning session, ROIs could be localized very consistently, with the mean distance between peak voxels in separate data sets being only 1.5 mm, accompanied by an average reduction in T-value of 1.6%. Between scanning sessions, distances between peak voxels were slightly higher (2.9 mm), and the reduction in T-value was greater (5.7%). Note that even across scanning sessions, the mean distance between peaks is less than the linear distance between adjacent voxels under typical scanning protocols. From these results we conclude that category-specific ROIs can be localized very consistently across runs and across sessions separated by weeks, thus providing a validation of the functional localization approach for higher order visual areas.

Why is between-session reproducibility worse than within-session reproducibility? This could be due to co-registration errors, attention and arousal differences, and/or priming effects.

The co-registration of functional to structural data is a likely contributor to the reduction in reproducibility between scanning sessions, as errors in co-registration will only affect between session comparisons. Automatic co-registration algorithms are now becoming available, and these may reduce the differences between within- and between-session reproducibility.

Another reason for lower reproducibility between sessions compared to within sessions could be a greater variability in attention and arousal across the two scanning sessions, resulting in greater variability of selectivity between sessions. Indeed, we found a greater absolute difference in T-values between sessions compared to within sessions.

Finally, it is possible that object-selective areas of the brain change over time, for example by forming sparser representations of stimuli after repetition (e.g., van Turennout et al., 2000). Repetition of stimuli can thus result in a reduction of brain activity to these stimuli, a phenomenon known as priming (see Henson, 2003 for a review). Although most studies reporting priming effects have used relatively short inter-stimulus intervals compared to the present study, it is possible that priming occurred even after three weeks. The finding that T-values remained nearly constant between scanning sessions, however, argues against priming as an explanation of the present results.

We found some evidence for differences in reproducibility between ROIs: between scanning sessions the PPA showed less reduction in T-value than the FFA and OFA. As the reliability of a finding is reflected in the statistical significance of this finding, we might expect higher T-values in the PPA than the FFA and OFA. Indeed, we found the highest T-values in the PPA, and the lowest T-values in the FFA and OFA. In the present study high T-values indicate a large difference in the response of a ROI to the

preferred category compared to the control categories, and/or small within-subject variance of this difference. Thus, it may be that perceptual and/ or cognitive differences between the preferred category and the remaining categories were lowest in the FFA and OFA. A recent finding showing strong body-selectivity in the fusiform gyrus close to the FFA, indicates that the representation of bodies and faces may partially overlap, suggesting that contrasts between faces and bodies may be less effective than between faces and other categories (Peelen and Downing, 2005a). More generally, choosing control categories that produce a low response in a ROI could enhance the within-subject reproducibility of the ROI.

To conclude, we find that category-specific visual areas can be localized very reliably within subjects, even across sessions separated by three weeks. This supports the validity of the functional localization approach for investigating these theoretically-important regions of extrastriate visual cortex. Future studies could use the methods adopted here to investigate the reproducibility of other areas. In particular, it would be interesting to compare the reproducibility of higher cognitive areas (e.g., memory related areas in prefrontal cortex) with lower level areas. Activations in higher level areas may be relatively variable between subjects, but equally reproducible within subjects (Miller et al., 2002; Noll et al., 1997).

Chapter 3:

Selectivity for the human body in the fusiform gyrus

ABSTRACT

Functional neuroimaging studies have revealed human brain regions, notably in the fusiform gyrus, that respond selectively to images of faces as opposed to other kinds of objects. Here we use fMRI to show that the mid-fusiform gyrus responds with nearly the same level of selectivity to images of human bodies without faces, relative to tools and scenes. In a group-average analysis (N=22), the fusiform activations identified by contrasting faces vs. tools and bodies vs. tools are very similar. Analyses of within-subjects regions of interest, however, show that the peaks of the two activations occupy close but distinct locations. In a second experiment, we find that the body-selective fusiform region, but not the face-selective region, responds more to stick figure depictions of bodies than to scrambled controls. This result further distinguishes the two foci, and confirms that the body-selective response generalises to abstract image formats. These results challenge accounts of the mid-fusiform gyrus that focus solely on faces, and suggest that this region contains multiple distinct category-selective neural representations.

INTRODUCTION

A major current theme in cognitive neuroscience is the effort to understand the brain systems involved in perceiving the identities, emotional states, and intentions of other people. To date, much of this research has focused on the perception of faces, with particular concentration on face-selective activity in the fusiform gyrus. Functional magnetic resonance imaging (fMRI) studies have shown that the "fusiform face area" (FFA) responds strongly and selectively to human faces as compared to a wide variety of controls (Halgren et al. 1999; Kanwisher et al. 1997; Puce et al. 1995). FFA activity closely tracks awareness of the presence of a face (Andrews et al. 2002; Andrews and Schluppeck 2004; Hasson et al. 2001; Tong et al. 1998), as well as trial-by-trial psychophysical performance on face detection and identification tasks (Grill-Spector et al. 2004). One interpretation of these findings is that the FFA represents a cortical "module" for face processing.

Strong selectivity for a stimulus class, however, does not necessarily imply a dedicated system for processing that class. One alternative proposal is that the FFA may instead be better conceived as a mechanism for distinguishing visually-similar exemplars of any object class for which the viewer has substantial expertise (Tarr and Gauthier 2000). This is supported, for example, by the finding that bird and car experts show increased FFA activity when viewing birds and cars, respectively (Gauthier et al. 2000). Others have gone further to suggest that the FFA, and highly-selective activation "peaks" in general, have no special functional role, but instead form part of a broad neural network that represents objects by a distributed collection of feature analysers (Haxby et al. 2001, but

see Spiridon and Kanwisher 2002). Thus the FFA has been at the centre of a broader debate about the organisation of ventral stream object representations (Cohen and Tong 2001; Kanwisher 2000; Levy et al. 2001).

The face, however, is not the only source of socially-relevant cues; the rest of the body also conveys such information. Although human bodies and faces are visually dissimilar, they have other features in common. Both provide cues to identity, emotion, intention, age, and gender. For both bodies and faces, the differences between exemplars are metric rather than qualitative, and for both most adults will have developed substantial perceptual expertise. Moreover, recent studies have shown behavioural effects for bodies that were generally thought to be specific for faces, such as a "body-inversion" effect (Reed et al. 2003) and an advantage for bodies in attentional capture (Downing et al. 2004).

There is some evidence that this conceptual and behavioural similarity between faces and bodies may be reflected in neural activity in the fusiform gyrus. One fMRI study found a significantly higher response in the FFA to bodies than to other object categories, although the response to bodies was significantly lower than to faces (Kanwisher et al. 1999). In this study, however, analysis of body-related activity was limited to a region functionally defined by contrasting faces vs. houses. Another fMRI study, in which subjects were scanned while viewing segments of a James Bond movie, suggested that bodies activate the fusiform gyrus near the FFA (Bartels and Zeki 2004). The stimuli in this study were not controlled, so it is not clear whether bodies, faces, or both were responsible for this activation. A recent ERP study showed that the face-selective N170 is also elicited by bodies (with faces blurred), suggestive of shared underlying neural

processes (Stekelenburg and de Gelder 2004). Note, however, that the relationship of the N170 to the FFA is unclear. Furthermore, the response to bodies in this study could also reflect contextual enhancement of the blurred faces (Cox et al. 2004). Finally, there is mixed evidence on whether the FFA response to the bodies of non-human animals is higher than the response to inanimate objects (Chao et al. 1999, Kanwisher et al. 1999, Grill-Spector et al. 2004).

Investigations of the macaque visual system also provide relevant findings. Single-unit recordings have revealed neurons distributed throughout the macaque temporal lobe, particularly in the superior temporal sulcus (STS), that respond to views of the face and body, and to actions involving them (Jellema and Perrett 2003; Perrett et al. 1982; Wachsmuth et al. 1994). More recently, fMRI in the macaque has shown face- and body-selective activations in adjacent regions of the lower bank of the STS (Tsao et al. 2003). A homology between this region and human FFA has been proposed (Gauthier and Logothetis 2000, Tsao et al. 2003) although conclusive evidence is still lacking. Thus, in the macaque the representations of faces and bodies appear to be intertwined, in a region that may map to the fusiform gyrus in humans.

The above evidence leads to the prediction that human bodies, even without faces, may be represented in the human mid-fusiform gyrus. This possibility has not been tested rigorously to date. Therefore the three objectives of the present study were: 1) to use fMRI to measure the selectivity of neural responses to headless bodies in the fusiform gyrus; 2) to compare the locations of body-related and face-related activity in this region; and 3) to test whether body selectivity in the fusiform gyrus generalises to more abstract depictions of the human form.

MATERIALS AND METHODS

Participants

Twenty-two healthy adult volunteers (12 female) were recruited from the University of Wales, Bangor community for the first experiment. Eight subjects (4 female; one of whom participated in the first experiment) participated in the second experiment. Participants satisfied all requirements in volunteer screening and gave informed consent. Ethics approval was obtained from the School of Psychology at the University of Wales, Bangor and the North-West Wales Health Trust. Participation was compensated at £20 per session.

Experimental paradigm

The first experiment consisted of images of faces, human bodies without heads, outdoor scenes, and handheld tools, presented in a blocked design (see Fig. 1A). There were 40 exemplars of each category, divided into two sets (A and B).

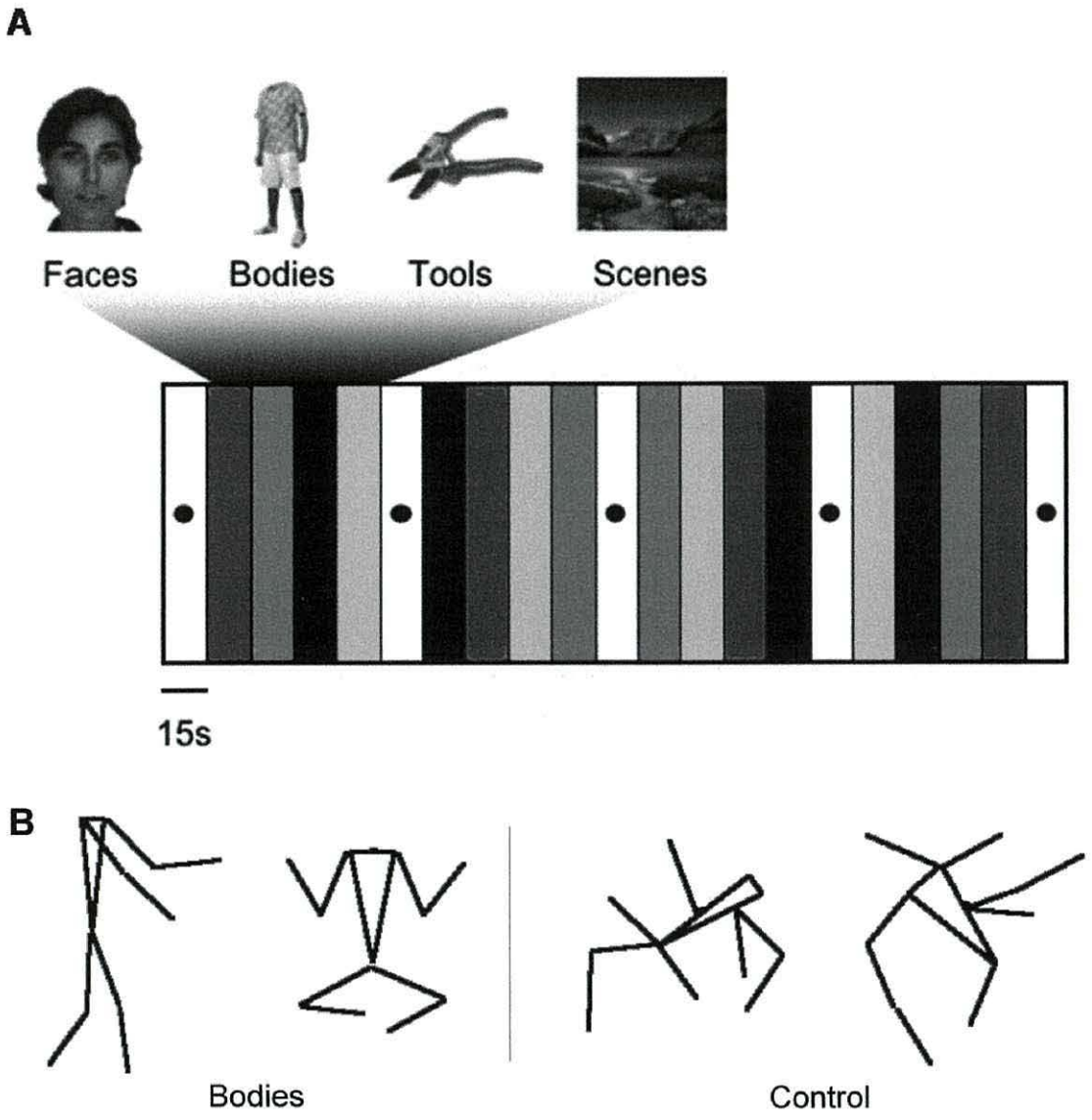


Figure 1. A) Schematic of the design of the first experiment, with sample stimuli. *B)* Sample stimuli from the second experiment.

Subjects performed 4 runs, each lasting 315 seconds (105 functional volumes). Runs 1 and 2 contained stimuli from set A, and runs 3 and 4 contained stimuli from set B. Each functional run consisted of 21 15-sec blocks. Blocks 1, 6, 11, 16, and 21 were a fixation-only baseline condition. Each of the remaining blocks comprised presentation of

20 exemplars from a single category. Within a block, each stimulus was presented for 300 msec, with an ISI of 450 msec, during which a central fixation point appeared on the screen. Two versions of block order were adopted. The first half and second half of one version were swapped to create the second version. The order of blocks was symmetrically counterbalanced within each version, so that the first half of each version was the mirror order of the second half. The result is that the mean serial position of each condition was equated, reducing the possible contribution of linear confounds to the results.

The second experiment tested 15 stick figures and 15 scrambled control items (see Fig. 1B). Fifteen-second blocks from each condition were alternated, with intervening fixation baseline blocks as above. Two orders of each design were tested between subjects to counterbalance for order effects. In each block, 15 images were presented (300 msec on, 700 msec off for each item).

In both experiments, subjects performed a "1-back" repetition-detection task, in which they were asked to press a button whenever an image occurred twice in immediate succession. Two image repetitions occurred at randomly-selected points in each block.

Functional imaging and analysis

Data acquisition. A 1.5T Philips MRI scanner with a SENSE (Pruessmann et al. 1999) parallel head coil was used. For functional imaging, an EPI sequence was used (TR = 3000 ms, TE 50 ms, flip angle 90°, FOV = 240, 30 axial slices, 64 x 64 inplane matrix, 5 mm slice thickness). The scanned area covered the whole cortex and most of the cerebellum.

Data analysis. Pre-processing and statistical analysis of MRI data was performed using BrainVoyager 4.9 (Brain Innovation, Maastricht, The Netherlands). The first three volumes of each run were discarded in order to avoid differences in T1 saturation. Functional data were motion-corrected and low-frequency drifts were removed with a temporal high-pass filter (0.006 Hz). Functional data were manually co-registered with 3D anatomical T1 scans (1 x 1 x 1.3 mm resolution), and then resampled to isometric 3 x 3 x 3 mm voxels with trilinear interpolation. The 3D scans were transformed into Talairach space (Talairach and Tournoux 1988), and the parameters for this transformation were subsequently applied to the co-registered functional data.

In order to generate predictors for the multiple-regression analyses, the event time series for each condition were convolved with a delayed gamma function (delta = 2.5 s; tau = 1.25 s) in order to model the hemodynamic response (Boynton et al. 1996). Voxel time series were *z*-normalized for each run, and additional predictors accounting for baseline differences between runs were included in the design matrix.

ROI definition and peak-voxel comparison. For each subject, we located the most significantly activated voxel for the contrast faces vs. tools and the contrast bodies vs. tools, within a restricted part of ventral cortex (Talairach coordinates: $30 < x < 50$, $-60 < y < -30$, $-30 < z < -5$). Regions of interest (ROIs) for the first experiment were defined as the set of contiguous voxels that were significantly ($p < 0.05$, uncorrected for multiple comparisons) activated within 8 mm in the anterior/posterior, superior/inferior, and medial/lateral direction of the most significantly activated voxel. For the first experiment, run 1 and 3 were combined to define the ROIs for runs 2 and 4, and vice versa. Thus the data used for ROI-definition were independent from the data reported. For the second

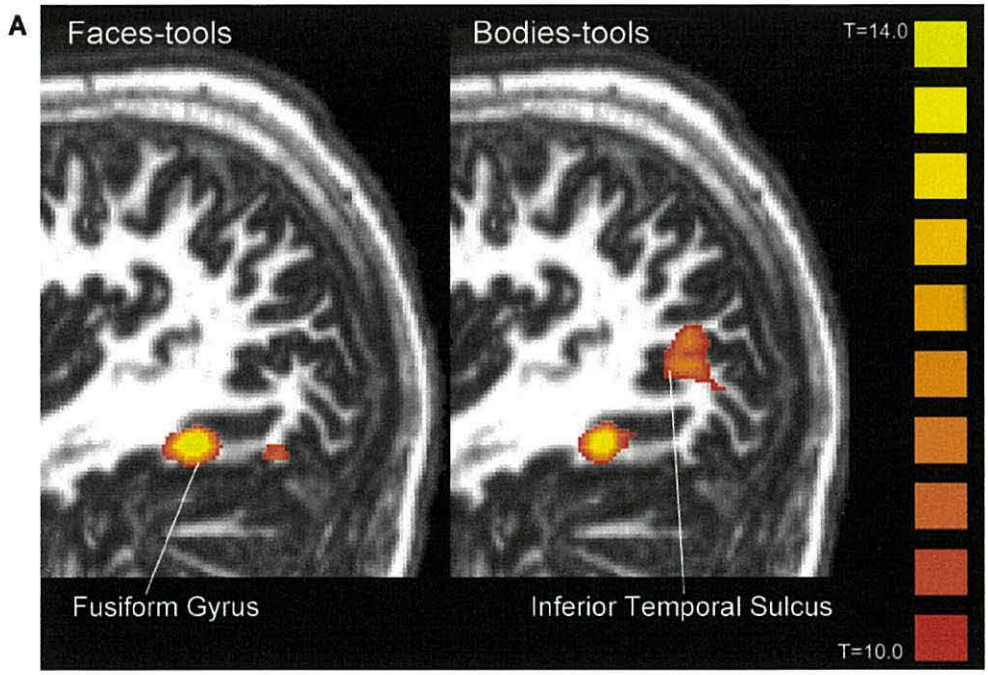
experiment, face- and body-selective ROIs were identified based on an independent set of localiser data.

To further examine the spatial relationship between ROIs, the locations of the most significantly activated voxels for faces and bodies (versus tools) were compared within subjects. We calculated three measures of the linear distance between the most significant voxels. Measure 1 was the distance between peaks within a category and stimulus set. The locations of the most significant voxels of run 1+3 and run 2+4 were compared, for faces and bodies separately. Thus, in this comparison, each set of data contained one run for each of the two stimulus sets. This measure serves as a baseline estimate of how the peak location for identical stimuli varies between scans. Measure 2 was the distance between peaks within a category but across stimulus sets. The locations of the most significant voxels of run 1+2 and run 3+4 were compared, again for faces and bodies separately. This tests how spatially consistent responses are to different exemplars of the same category. Measure 3 was the distance between peaks across categories. The locations of the most significant voxels of run 1+3 (e.g. for faces) and run 2+4 (e.g. for bodies) were compared. Likewise, peak voxels from run 1+2 and run 3+4 were compared, again across categories. This measure tests whether the face- and body-selective foci occupied different locations.

Note that in these measures, the *mean* distance between peak voxels can be smaller than the distance between two adjacent voxels in a single scan. These distance measures are conservative in the sense that they could fail to detect differences in the location of peaks that are smaller than the scanned resolution, but could not artifactually create such a difference where one did not exist at a finer scale.

RESULTS

A whole-brain group-averaged multiple regression analysis (fixed effects) contrasting faces vs. tools and bodies vs. tools revealed significant activation in the right fusiform gyrus for both contrasts. The fusiform activations were comparable in location and extent (Fig. 2A). The face contrast, but not the body contrast, activated a right ventral occipital region, corresponding to the occipital face area (OFA [Puce et al. 1996]). The body contrast, but not the face contrast, activated an occipito-temporal region bilaterally, corresponding to the "extrastriate body area" (EBA [Downing et al. 2001]).



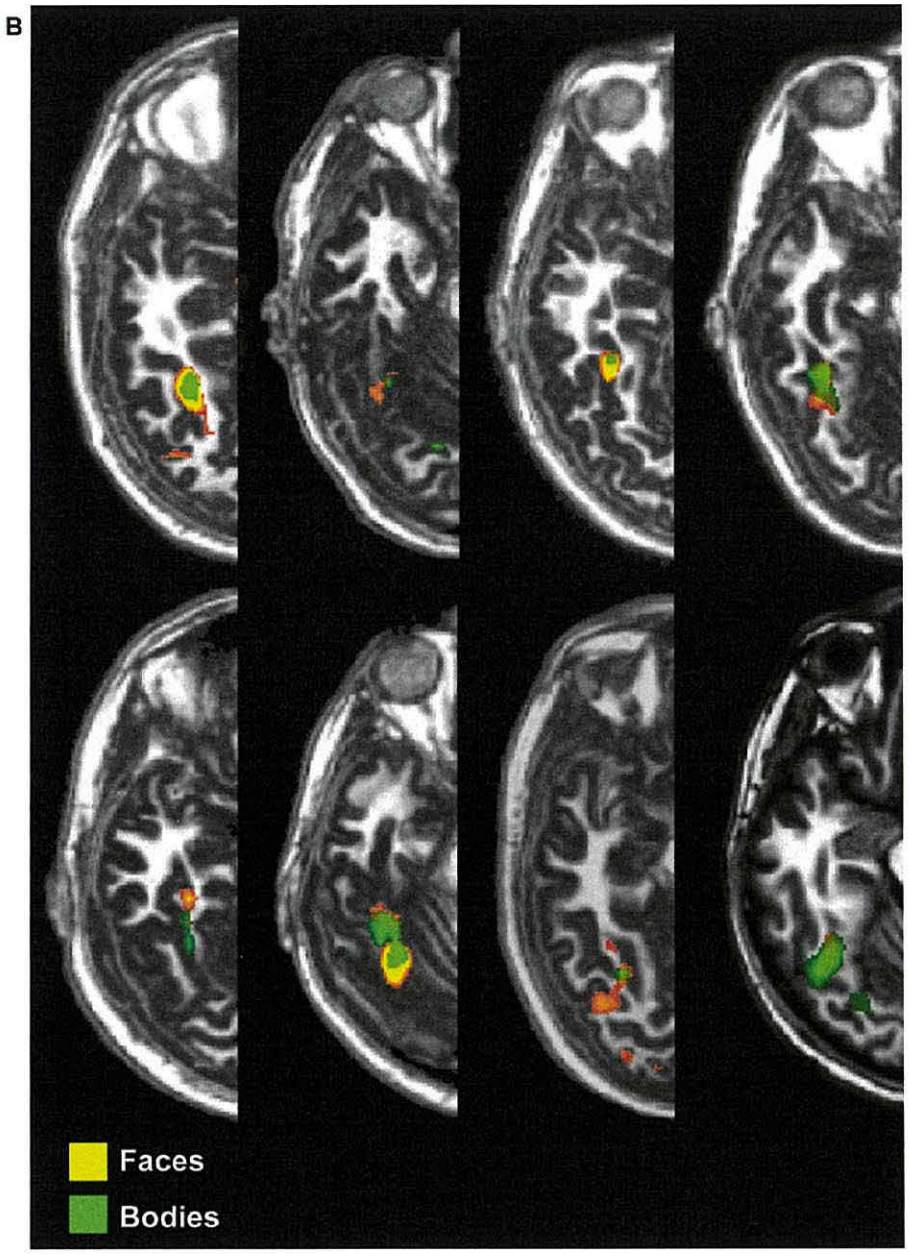
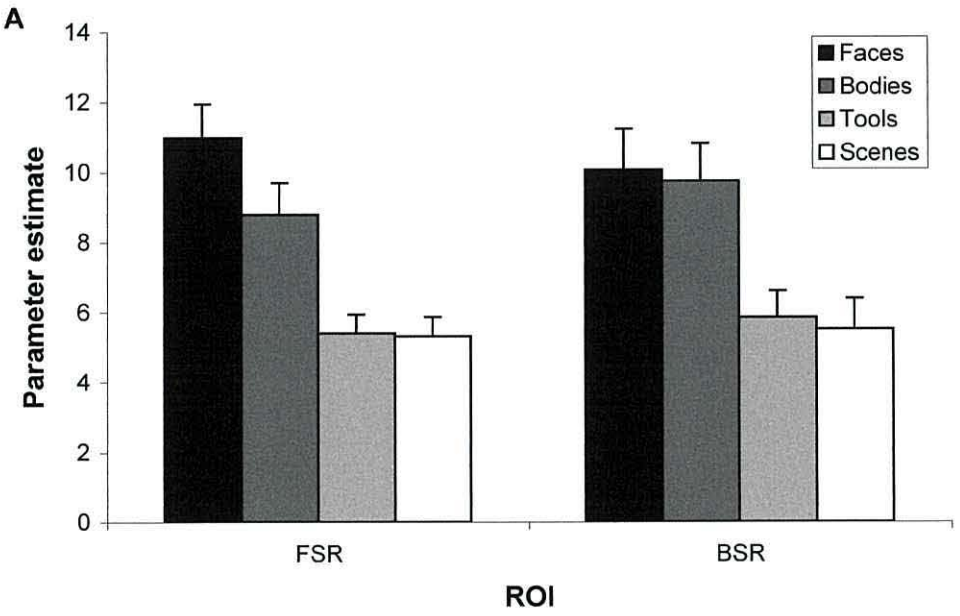


Figure 2. A) Group activation maps (T > 10) depicted on a sagittal slice (x = 39). The contrast faces vs. tools (left panel) showed significant activation in fusiform gyrus (FFA; Talairach coordinates: 39, -44, -18; volume: 0.8 cm^3) and ventral occipital cortex (OFA; 39, -64, -20; 0.1 cm^3). The contrast bodies vs. tools (right panel) showed significant activation in fusiform gyrus (40, -43, -17; 0.8 cm^3) and inferior temporal sulcus (EBA; 45, -65, 2; 4.8 cm^3). *B)* Activation maps (T > 5) showing face-selective and body-selective fusiform activations in eight individual subjects. Orange: faces vs. tools; Green: bodies vs. tools.

For each subject, we identified the right hemisphere fusiform face-selective region (FSR; defined by faces vs. tools) and the right hemisphere fusiform body-selective region (BSR; defined by bodies vs. tools) as regions of interest (Fig. 2B and Table 1). (The labels we give these regions are intended for explanatory convenience; we do not necessarily assume that there are two functionally-independent regions). For each ROI, we calculated the mean parameter estimate from the regression analysis for bodies, faces, tools and scenes (Fig. 3A).



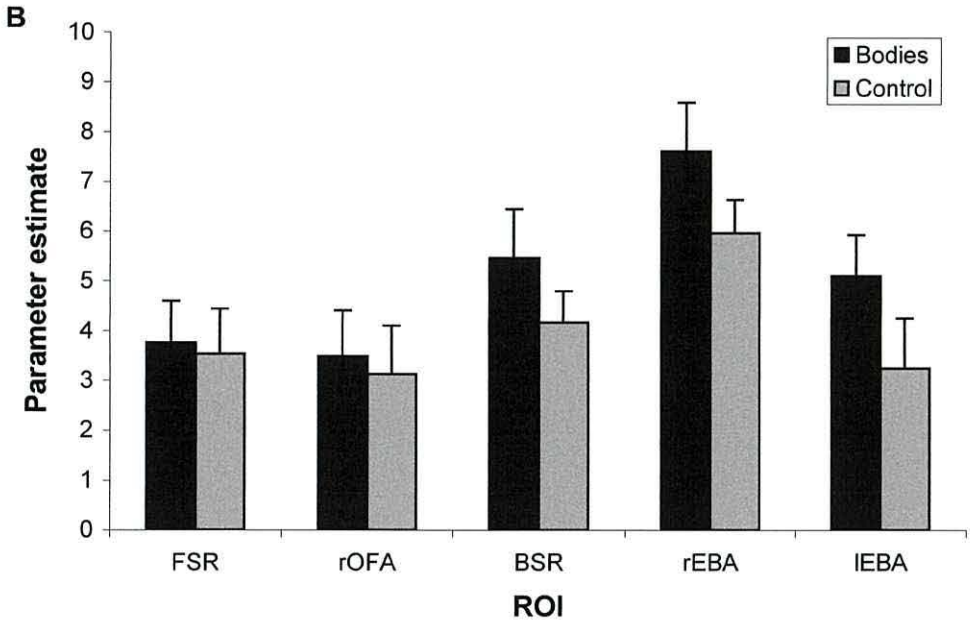


Figure 3. A) Parameter estimates for faces, bodies, tools and scenes in the face-selective (FSR) and body-selective (BSR) fusiform regions of interest (as defined by faces vs. tools and bodies vs. tools respectively). ROIs were defined with independent data sets. Selectivity in the face-selective region (ratio to tools) was 2.0 for faces and 1.6 for bodies. Selectivity in the body-selective region was 1.7 for both faces and bodies. Error bars indicate standard error of the mean. *B)* Parameter estimates for body stick figures and scrambled controls in key regions of interest. Error bars indicate standard error of the mean. FSR: fusiform face-selective region; OFA: occipital face area; BSR: fusiform body-selective region; EBA: extrastriate body area.

Table 1. Talairach coordinates and T values of the most significant voxel for faces and bodies (versus tools) in the mid-fusiform gyrus.

	Faces-Tools				Bodies-Tools			
Subject	**x**	**y**	**z**	***T***	**x**	**y**	**z**	***T***
1	38	-50	-15	9.0	44	-38	-18	4.9
2	38	-47	-15	9.7	41	-41	-15	7.4
3	35	-38	-24	8.6	35	-47	-24	6.8
4	35	-53	-18	7.4	35	-47	-21	3.9
5	35	-53	-18	18.8	38	-41	-21	9.8
6	35	-47	-15	15.9	35	-47	-12	9.5
7	44	-50	-15	6.7	44	-50	-15	7.8
8	44	-47	-15	8.9	44	-44	-15	5.6
9	44	-38	-12	4.2	38	-41	-18	4.2
10	38	-44	-21	9.7	41	-44	-18	9.3
11	38	-53	-12	13.0	41	-44	-15	8.6
12	35	-50	-12	6.2	38	-56	-15	10.3
13	41	-53	-24	6.9	38	-47	-24	6.5
14	35	-32	-15	9.1	35	-41	-9	9.0
15	38	-44	-24	9.8	44	-38	-21	6.7
16	38	-38	-12	9.2	35	-38	-15	5.7
17	41	-32	-15	5.6	38	-44	-15	5.0
18	41	-56	-24	8.5	41	-50	-24	4.4
19	44	-38	-21	7.7	44	-38	-21	5.5
20	38	-35	-21	8.6	44	-41	-21	3.4
21	35	-47	-18	15.1	35	-47	-18	10.5
22	35	-53	-27	9.7	35	-41	-27	8.0

A repeated-measures ANOVA revealed a significant interaction between category and ROI ($F_{3,19} = 7.4$, $p < 0.005$). The FSR responded more to faces than bodies ($t_{21} = 4.2$, $p < 0.001$), tools ($t_{21} = 9.7$, $p < 0.001$) and scenes ($t_{21} = 8.1$, $p < 0.001$), and more to bodies than tools ($t_{21} = 7.3$, $p < 0.001$) and scenes ($t_{21} = 6.3$, $p < 0.001$), consistent with previous findings in the FFA (Kanwisher et al. 1999). In contrast, the BSR did not distinguish faces and bodies ($t_{21} = 0.5$, $p = 0.60$), and responded more to faces and bodies than tools ($t_{21} = 6.9$, $p < 0.001$ and $t_{21} = 7.4$, $p < 0.001$) and scenes ($t_{21} = 5.9$, $p < 0.001$ and $t_{21} = 7.8$, $p < 0.001$). No differences were observed in the response to tools and scenes

in either region (both $ps > 0.35$). The interaction between category and ROI indicates that the two ROIs occupied at least partially different locations.

To compare the location of fusiform body-related and face-related activity within subjects, we calculated the distance between the most significant voxels: (1) within-category, within-stimulus set; (2) within-category, between-stimulus set; and (3) between-categories (see methods). No effect of stimulus set (measure 1 vs. measure 2) was observed ($F_{1,21} = 0.02$, $p = 0.9$), nor did this depend on category ($F_{1,21} = 0.4$, p=0.54), showing that different exemplars of the same category produce similar activation peaks. In contrast, we found a significantly larger distance across categories (measure 3; 7.5 mm, SE = 0.7) than within categories (measure 2; 5.7 mm, SE = 0.6), $t_{21} = 3.0$, $p < 0.01$, indicating distinct peaks of activation for bodies and faces. We also tested whether there was a consistent spatial relationship between the FSR and BSR within subjects. The difference in locations between ROIs did not significantly differ from zero in the x, y, and z directions, all $p > 0.20$, indicating that there was no consistent relationship between the two ROIs, at least in 3D space.

In a second experiment, we compared headless stick figure depictions of bodies and scrambled control figures (Fig. 1B). Abstract stimuli such as these minimize the differences in image features between body and non-body stimuli. A significant interaction, $F_{1,7} = 7.6$, $p < 0.05$, showed that the difference between stick figures and controls differed between the BSR and the FSR. The response to stick figures was significantly higher than to controls in the BSR ($t_7 = 2.6$, $p < 0.05$), but not in the FSR ($t_7 = 0.5$, $p = 0.62$). The OFA likewise showed no significant difference, ($t_7 = 1.1$, $p = 0.33$). There was a significant effect in both the right EBA ($t_7 = 3.4$, $p < 0.05$) and the left EBA

(t_7=3.5, p<0.05), replicating previous findings (Downing et al. 2001). Response magnitudes in all of these regions are given in Figure 3B. The selective response to stick figures as compared to scrambled controls was strong enough to show up in a whole-brain, group-averaged, fixed-effects multiple regression analysis. Significant activation (p < 0.0005 uncorrected) was found in the right fusiform gyrus (peak Talairach coordinates: 41, -38, -21) and bilaterally in the vicinity of the EBA (left: -49, -74, 9; right: 47, -62, 6).

DISCUSSION

We found nearly identical activation in the fusiform gyrus for faces and bodies (vs. tools) in a whole-brain group analysis. Analyses of individually-defined ROIs, however, showed a different selectivity pattern for the region defined by faces compared to that defined by bodies. There were distinct peaks of activation within the fusiform gyrus for the two different categories, but not for different stimulus sets within a category. Finally, we confirmed that the body-selective fusiform region (but not the face-selective region) distinguished intact and scrambled versions of human stick-figures. This result shows that the selective response to bodies found here is not entirely due to differences in low-level image features between bodies and other objects.

One account of our findings could be that perceiving bodies leads observers to mentally image the missing faces, which in turn activates face-selective neurons. This is unlikely to explain our results, for several reasons. The body stimuli were unfamiliar and not associated to a particular face; it may not be possible to mentally image a "generic" face in this situation. Second, a previous study has compared the FFA response to front,

side, and back views of heads (Tong et al. 2000). The back of the head, particularly when viewed in context with frontal views of the same individuals, would be expected to induce mental imagery for the missing face at least as strongly as bodies, yet FFA responses to this condition were comparable to inanimate objects. Finally, even explicit instruction to image specific faces produces relatively weak FFA activation that is not sufficient to explain the strong body selectivity seen here (O'Craven and Kanwisher 2000).

A related proposal is that bodies provide a context that causes ambiguous perceptual input to be perceived as a face. Cox et al. (2004) showed that the mid-fusiform activity produced by a gaussian blur was increased when it was shown in a face-like contextual relationship to a body (i.e., atop the neck). (The selectivity of this region for faces:scenes, however, was low, raising the possibility that it did not correspond perfectly to the FFA as defined in previous reports). Here, however, we showed only headless bodies without any stimulus that could be contextually enhanced.

Another possible account of the present findings relates to retinal eccentricity (Levy et al. 2001). Lateral visual areas, including the fusiform gyrus, are thought to be important for center-biased visual field representations, whereas more medial areas are thought to be important for periphery-biased field representations. Bodies and faces are likely to be represented more centrally than scenes, which could account for the higher response to bodies and faces relative to scenes. Tools, however, often contain fine details that need central processing to discriminate, and the response in this region to tools was relatively low. Moreover, we found no significant difference in either fusiform ROI

between the response to tools and scenes, arguing against a central visual field bias in this part of the fusiform gyrus.

How do the present findings improve our understanding of the ventral temporal lobe? The mid-fusiform gyrus is part of a non-retinotopic, high-level object representation system. Our results show that some neurons in this region distinguish bodies, as well as faces, from other object kinds. Other object categories (such as cars, birds, and "greebles") have also been shown to produce an enhanced, though much less selective, response in this region, particularly in subjects who are experts at distinguishing among exemplars of these categories (Gauthier et al. 2000). One possibility is that the same population of neurons is engaged by all of these categories. Activity in these neurons could thus reflect cognitive processes that are independent of the stimulus categories involved, such as distinguishing among visually similar exemplars. This seems improbable, at least for the case of faces and bodies, however, in light of our finding that selective foci for these categories occupy distinct locations. We suggest instead that different, possibly overlapping, populations of selective neurons co-exist at close quarters. Faces and bodies are represented by larger populations and/or more selective neurons than other categories. Existing data do not speak to the root cause of this bias, which could be present from birth or could develop through experience.

These findings must also be considered in light of the recent proposal that objects are represented in the ventral stream by distributed patterns of neural activity. On this view, highly-selective peaks, such as those seen here for faces and bodies, play no unique role in this representation (Haxby et al. 2001). One way to test this prediction is to ask for which brain regions activity is systematically related to performance on perceptual tasks.

A recent study using this approach (Grill-Spector et al. 2004) showed a positive correlation between successful face detection and identification and FFA activity. This relationship was not observed in other ventral areas, even those (such as the "parahippocampal place area"; Epstein and Kanwisher 1998) where the response to faces is consistently low and could in principle provide information relevant to face processing. This finding suggests that regions with high category selectivity may indeed have a privileged role in vision.

These considerations lead to possible avenues for further research. Scanning at higher resolution may help to elucidate the spatial relationship among sub-regions of the fusiform gyrus that show selectivity for faces, bodies, or other object kinds. Further stimulus or task manipulations will be necessary to determine whether these selective foci can be functionally dissociated. Finally, it will be important to determine for which categories activity in the fusiform region predicts trial-by-trial performance on perceptual tasks. If this relationship is found for the body-selective region identified here, it would provide additional evidence that the fusiform gyrus plays a previously unsuspected role in visual processing of the human body.

Chapter 4:

Patterns of fMRI activity dissociate overlapping functional brain areas that respond to biological motion

ABSTRACT

Accurate perception of the actions and intentions of other people is essential for successful interactions in a social environment. Several cortical areas that support this process respond selectively in fMRI to static and dynamic displays of human bodies and faces. Here we apply pattern-analysis techniques to arrive at a new understanding of the neural response to biological motion. Functionally defined body-, face-, and motion-selective visual areas all responded significantly to "point-light" human motion. Strikingly, however, only body selectivity was correlated, on a voxel-by-voxel basis, with biological motion selectivity. We conclude that: a) biological motion, through the process of structure-from-motion, engages areas involved in the analysis of the static human form; b) body-selective regions in posterior fusiform gyrus and posterior inferior temporal sulcus overlap with, but are distinct from, face- and motion-selective regions; c) the interpretation of region-of-interest findings may be substantially altered when multiple patterns of selectivity are considered.

INTRODUCTION

One of the most important functions of vision is to provide information about the actions, intentions, and identities of other individuals. Functional magnetic resonance imaging (fMRI) research into how the human visual cortex accomplishes this task has identified neural activity in a number of posterior areas that are selective for the visual appearance of conspecifics. In the human, the extrastriate body area (EBA; Downing et al., 2001b), which is found at the posterior end of the inferior temporal sulcus (partly overlapping motion-selective area hMT+), is selective for static images of human bodies and body parts. Face-selective responses are found in posterior fusiform gyrus, particularly in the right hemisphere (fusiform face area or FFA; Kanwisher et al., 1997). More recently, strongly body-selective responses have also been reported in the posterior fusiform gyrus, in a region closely overlapping the FFA: the "fusiform body area" or FBA (Peelen and Downing, 2005a; Schwarzlose et al., 2005). Finally, realistic and schematic biological movements of the hands, face, and whole body reliably activate posterior superior temporal sulcus (pSTS; Allison et al., 2000; Beauchamp et al., 2002; Bonda et al., 1996; Haxby et al., 2002; Pelphrey et al., 2005).

A major focus in the neuroimaging research on biological motion processing has been on selective neural responses to "point-light" (PL) animations (Grezes et al., 2001; Grossman et al., 2000, 2004; Grossman and Blake, 2001, 2002; Michels et al., 2005; Peuskens et al., 2005; Santi et al., 2003; Saygin et al., 2004; Vaina et al., 2001). These displays, originally conceived by Johansson (1973), consist of only a few dots that move in a way characteristic of human movements, e.g. walking or jumping. They are of

theoretical interest because they convey biological motion patterns with little or no form information in individual frames, and without most of the visual features (e.g., clothes, skin) normally present in images of moving human bodies. Interpretation of fMRI studies comparing PL figures to various controls has largely focused on pSTS. But biological motion consistently activates other posterior regions, even when low-level factors such as the presence of visual motion *per se* are accounted for. To date, the presence of these activations and their functional significance remains unexplained.

Several studies have reported activations to PL action animations in fusiform gyrus (Grossman et al., 2002, 2004; Santi et al., 2003). Grossman and colleagues (2004) have interpreted this finding as reflecting engagement of face-selective FFA by PL animations. However, given its close proximity, this activation could instead (or additionally) reflect activation of body-selective FBA. PL figures also activate the posterior inferior temporal sulcus / middle temporal gyrus (Michels et al., 2005; Peuskens et al., 2005; Saygin et al., 2004). It is at present unclear whether this activation reflects activation of body-selective neurons in the EBA (compare Downing et al., 2001b and Michels et al., 2005, with Grossman and Blake, 2002), or instead, motion-selective neurons in area hMT+.

Here we provide a simplifying resolution to these open questions, by demonstrating that PL-related responses outside of pSTS reflect engagement of known neural regions that are selective for static images of the human body. By performing voxel-by-voxel analyses of the response patterns to different stimuli, we were able to disentangle body-selective from face- and motion-selective responses in posterior fusiform gyrus and inferior temporal sulcus. In all regions tested, body selectivity, but not

face or motion selectivity, could predict the response to biological motion displays on a voxel-by-voxel basis.

MATERIALS AND METHODS

Subjects

Eighteen healthy adult volunteers were recruited from the University of Wales, Bangor community. Participants satisfied all requirements in volunteer screening and gave informed consent approved by the School of Psychology at the University of Wales, Bangor and the North-West Wales Health Trust. Participation was compensated at £20 per session.

Design and Procedure

Each participant was scanned on three blocked-design fMRI experiments, in order to identify *a priori* functional regions of interest with respect to individual brain anatomy, and to measure the response of these regions to biological motion stimuli. We localised the EBA and FBA with an experiment consisting of blocks of images of human faces, human bodies without heads, outdoor scenes, and handheld tools. The experiment consisted of 21 15-sec blocks. Blocks 1, 6, 11, 16, and 21 were fixation-only baseline epochs. In each of the remaining blocks, 20 different images from one category were presented. Each image appeared for 300 msec, followed by a blank screen for 450 msec. Twice during each block, the same image was presented two times in succession. Subjects were required to detect these repetitions and report them with a button press (1-

back task). Image position was jittered slightly on alternate presentations, in order to disrupt attempts to perform the 1-back task based on low-level visual transients. Each participant was tested with two different order versions of the experiment, counterbalancing for the order of the blocks. In both versions, assignment of category to block was counterbalanced, so that the mean serial position in the scan of each condition was equated. Participants were tested with two (N=14) or four (N=4) runs of this experiment. Further details can be found elsewhere (Peelen and Downing, 2005a).

The blocked-design localiser for area hMT+ consisted of a pattern of low-contrast, concentric rings that either slowly oscillated inwards and outwards, or, in separate blocks, remained static (cf. Tootell et al., 1995). The experiment consisted of 21 15-sec blocks. Blocks 1, 6, 11, 16, and 21 were fixation-only baseline epochs. In the remaining blocks, moving and static stimuli were alternated. The stimuli in this experiment were passively viewed. Participants were tested with one run of this experiment.

The biological motion experiment had a similar design to the hMT+ localiser, except that the blocks were 16 sec long. In non-baseline blocks, subjects passively viewed either 16 intact point-light animations of simple, whole-body actions (e.g. jumping, or throwing), or scrambled controls of the same animations. Scrambled controls were made by keeping the component motions intact while randomizing the starting point of each dot. The animations were comprised of small white dots on a black background. Each animation lasted 667 msec, with a 333 msec blank interval before the next stimulus. Participants were tested with one (N=13) or two (N=5) runs of this experiment.

Data Acquisition

A 1.5T Philips MRI scanner with a SENSE parallel head coil was used. For functional imaging, a single-shot EPI sequence was used (T2* weighted, gradient echo sequence, TE = 50 ms, flip angle 90°). Scanning parameters were: TR=3000 ms, 20-22 off-axial slices, voxel dimensions: 3.75 x 3.75 x 5 mm (N=11), 3 x 3 x 4 mm (N=6), or 2.5 x 2.5 x 2.5 mm (N=1).

Pre-processing

Pre-processing and statistical analysis of MRI data was performed using BrainVoyager 4.9 (Brain Innovation, Maastricht, The Netherlands). Three dummy volumes were acquired before each scan in order to reduce possible effects of T1 saturation. Functional data were motion-corrected, low-frequency drifts were removed with a temporal high-pass filter (0.006 Hz). No spatial smoothing was applied. Functional data were manually co-registered with 3D anatomical T1 scans (1 x 1 x 1.3 mm resolution). The 3D anatomical scans were transformed into Talairach space, and the parameters for this transformation were subsequently applied to the co-registered functional data, which were resampled to 1 x 1 x 1 mm voxels. The analyses were performed in Talairach space to allow comparison of ROI locations with previous (and future) studies. Because the normalization parameters were (for a given subject) identical for all conditions, normalization could not have systematically influenced the results (cf. Swallow et al., 2003).

Whole-brain analysis

A whole-brain, random-effects group average analysis was conducted on data from the biological motion experiment. Seventeen subjects were included in this analysis (one subject was excluded because we did not scan the whole brain in this subject). A contrast was performed at an uncorrected threshold of $p < 0.001$ to test for regions more active in the intact than the scrambled conditions. Only clusters > 100 mm^3 are reported for this analysis.

ROI Analysis

For each participant, general linear models were created for each localiser experiment. One predictor (convolved with a standard model of the HRF) modeled each condition. Regressors of no interest were also included to account for differences in the mean MR signal across scans. Regressors were fit to the MR timeseries in each voxel, and the resulting parameter estimates were used to estimate the magnitude of response to each experimental condition.

In each participant, the localiser scans were used to define the EBA by contrasting the response to human bodies with that to the average of faces, tools, and scenes. The FBA was defined by contrasting bodies against tools, and the FFA was defined by contrasting faces against tools (Peelen and Downing, 2005a). (Note that in a recent study (Downing et al., in press) we found that identification and characterisation of these regions was little affected by the choice of baseline conditions.) We identified area hMT+ by contrasting the response to moving concentric rings with that to static rings. Analyses of the FBA and FFA were restricted to right hemisphere ROIs, on the basis of previous

evidence these regions are weaker or non-existent in the left hemisphere (Kanwisher et al., 1997; Peelen and Downing, 2005a).

For each ROI in each subject, the most significantly activated voxel was identified within a restricted part of cortex based on previously-reported anatomical locations (EBA: Peelen and Downing, 2005b; hMT+: Dumoulin, et al., 2000; FFA: Kanwisher et al., 1997; FBA: Peelen and Downing, 2005a). ROIs were defined as the set of contiguous voxels that were significantly activated (all $p < 0.001$ uncorrected) within a 9x9x9 mm cube surrounding (and including) the peak voxel. This procedure was adopted for four reasons: to ensure that regions were defined objectively, to ensure that they were segregated from nearby selective activations, to roughly equate the number of voxels included across different regions of interest, and to ensure that only the most selective voxels were included in the ROI. Within each ROI in each subject, a further general linear model was then applied, modeling the response of the voxels in the region (in aggregate) to the intact and scrambled biological motion conditions. The regression weights from this GLM provided the basis for the ROI amplitude results shown in Figure 2. A further analysis testing only the peak voxel of each ROI confirmed these results.

Correlation analyses

For each ROI in each subject individually, we measured the voxel-by-voxel pattern of selectivity to key stimuli of interest. This was accomplished by extracting a t-value for a given contrast at each voxel in the ROI. The t-value provides a useful index of selectivity, because it combines in one number the magnitude of the difference between two conditions, relative to the within-condition variance. For all regions of interest,

selectivity was measured for intact compared to scrambled point-light biological motion. For the pITS ROIs (hMT+ and EBA) motion selectivity was measured with moving vs static rings, and body selectivity with bodies vs the average of faces, tools, and scenes. For the pFG ROIs (FFA and FBA), face and body selectivity was measured with faces vs tools and with bodies vs tools, respectively.

These t-values provided the raw materials for further analyses of the relationship between different kinds of selectivity within a given ROI. For the first type of analysis, we correlated, for each ROI, the pattern of selectivity for one contrast, with the pattern for another. Thus, for example, we might correlate body selectivity and biological motion selectivity within the right EBA of a particular subject. These correlations were extracted for each subject individually, Fisher transformed, and the resulting mean correlation was tested statistically against zero.

For the second type of analysis, we computed similar correlations for ROIs defined by the union of the voxels of two individual ROIs (e.g., the union of right hMT+ and right EBA). This provides a test of the overall relationship between types of selectivity within a general region (e.g. posterior fusiform gyrus in the case of FFA/FBA; posterior ITS in the case of EBA/hMT+).

Finally, we used multiple regression to analyse how well biological motion selectivity could be predicted within a given ROI based on one kind of selectivity, while simultaneously taking into account another type of selectivity. For example, this analysis allowed us to ask to what extent biological motion selectivity could be predicted in EBA as a function of body selectivity, while accounting for any variance explained by motion selectivity. For each ROI in each subject, a linear model was fit with two kinds of

selectivity as predictors, and biological motion selectivity as the to-be-predicted variable. An additional "flat" predictor was included to account for global differences in selectivity. The fit from each model resulted in a normalized beta value for each of the two predictors of interest. These betas were collected from each subject individually, and compared against zero with a t-test.

RESULTS

Our experimental approach was as follows. First, we used a whole-brain, group-average analysis in order to identify gross regions that respond more to human actions, rendered as point-light animations, than to scrambled controls. Second, we identified several functional ROIs: hMT+, EBA, FFA, and FBA. We then measured the response of these individually-defined ROIs to the biological motion stimuli. Finally, we performed a series of voxel-by-voxel pattern analyses (Cox and Savoy, 2003; Haxby et al., 2001; Haynes and Rees, 2005; Kamitani and Tong, 2005) on individually-defined ROIs, with the goal of discovering the relationship between biological motion selectivity and motion, face, and body selectivity in those regions.

Eighteen subjects were tested on three blocked-design experiments. In the main experiment, point-light renderings of simple whole-body actions were compared to scrambled versions of the same sequences. As in a previous study (Grossman et al., 2000) the scrambled condition was created by randomizing the starting points of the light points from the intact sequences, but keeping the motion patterns intact. The other two experiments were used to identify functional regions of interest in each subject. One

compared oscillating to static low-contrast rings, in order to identify hMT+. The other experiment compared the responses to bodies (without heads), faces, scenes, and tools, in order to localize the body-selective regions EBA and FBA, and face-selective FFA.

Whole-brain analysis

An initial whole-brain group-average contrast between the point-light biological motion display and the scrambled control motion display revealed activation in various visual areas (Table 1). Replicating previous studies, strong activation was found in right posterior superior temporal sulcus (pSTS). We also found significant activation in bilateral posterior inferior temporal sulcus (pITS) and posterior fusiform gyrus (pFG). The peak coordinates of the pITS activation were close to those typical both for the EBA (Downing et al., 2001b; Peelen and Downing, 2005b) and for hMT+ (Dumoulin et al., 2000). The right fusiform activation fell close to the typical coordinates of the FFA and the FBA (Kanwisher et al., 1997; Peelen and Downing, 2005a). Thus the group-average analysis confirms the existence of significant biological motion selectivity in pITS and in pFG. The aim of the following analyses was to examine this selectivity in closer detail, on an individual subject level, in order to determine its source.

Table 1. Group-average activations for which biological motion > scrambled motion, from a random-effects multiple regression analysis, thresholded at $p < 0.001$ (uncorrected for multiple comparisons) and a minimum cluster size of 100 mm^3. Each row gives the anatomical location of the activation, the Talairach coordinates of the peak voxel, the mean T value, and the volume of activation. ITS = inferior temporal sulcus, STS = superior temporal sulcus.

	Talairach Coordinates				
Region	X	Y	Z	Mean T	mm^3
R. ITS	44	-69	-7	4.63	2101
R. ITS	52	-51	3	4.49	920
R. STS	57	-41	21	4.17	112
R. Fusiform	36	-39	-19	4.72	748
R. Fusiform	34	-68	-18	4.32	174
R. Post. Occipital	14	-96	-4	4.43	151
L.ITS	-42	-70	-4	4.59	2544
L. Supramarginal	-56	-39	25	4.37	388

Regions of interest

The EBA, hMT+, FBA, and FFA were localised in each subject individually. The mean sizes of the ROIs were (mm3 [SD]): left EBA (586[146]), right EBA (605[114]), left hMT+ (595[97]), right hMT+ (554[168]), right FBA (193[145]), right FFA (323[167]). Average peak Talairach and Tournoux (1988) coordinates for the EBA and hMT+ were (x [SD], y [SD], z [SD]): left EBA (-45[5], -74[4],-1[8]), right EBA (48[5], -70[5], 1[6]), left hMT+ (-44[5], -68[4], -4[8]), right hMT+ (44[4], -66[6], -2[7]). Average peak Talairach coordinates for the individually-localized right hemisphere FBA and right FFA were: FBA (41[3], -45[7], -19[5]), FFA (39[4], -47[7], -19[4]). Note that the average coordinates of EBA and hMT+ (within each hemisphere) were very similar, and that both were equally close to the pITS activation to biological motion in the group average analysis (left: -42, -70, -4; right: 44, -69, -7). Furthermore, even within single subjects,

EBA and hMT+ overlapped substantially (Figure 1a). Likewise, the average coordinates of FBA and FFA were similar to each other, and were similar to the group average biological motion fusiform gyrus activation (36, -39, -19). These regions also overlapped substantially within subjects (Figure 1b).

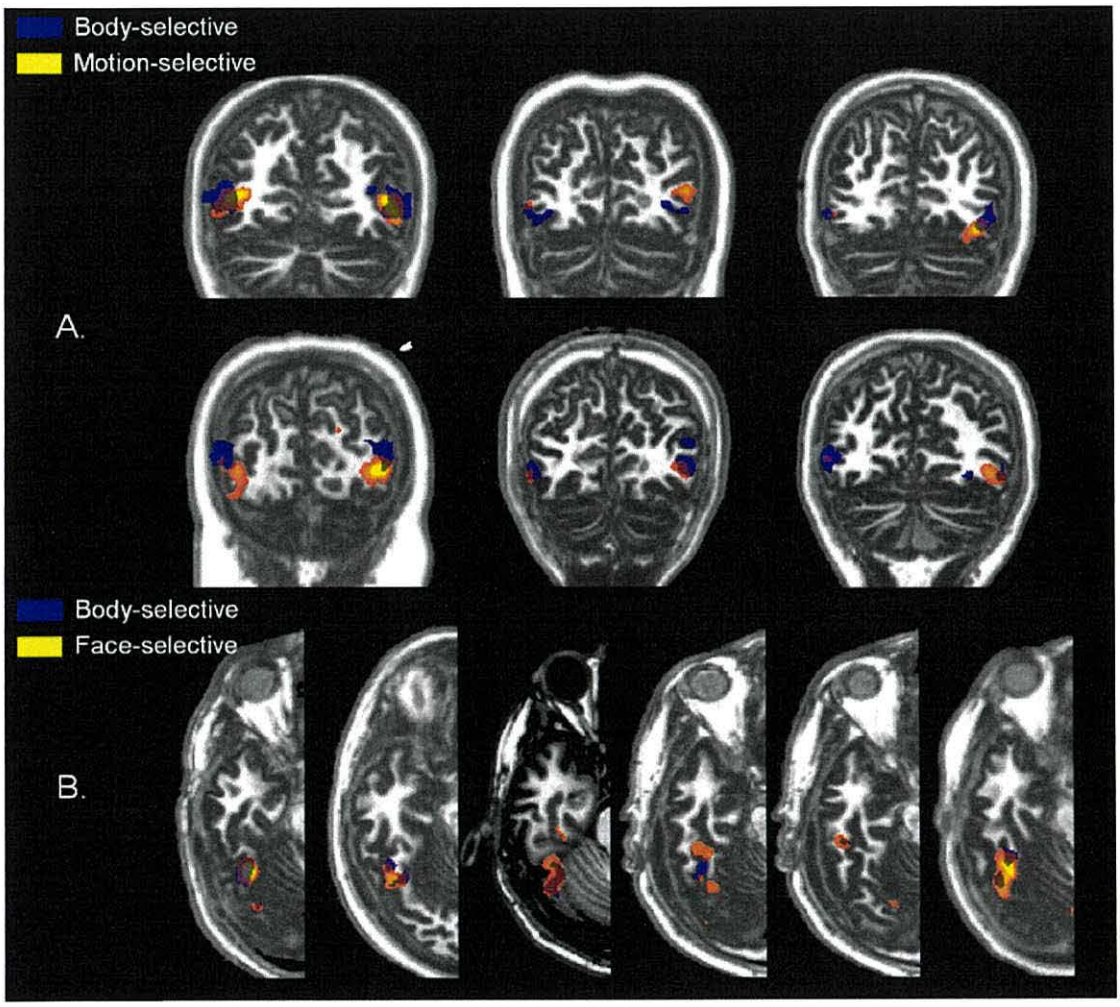


Figure 1. Overlap of nearby brain areas in individual subjects

(A) Body-selective (blue) and motion-selective (yellow) activations in posterior ITS in six individual subjects, at $p < 0.001$. (B) Body-selective (blue) and face-selective (yellow) activations in posterior fusiform gyrus in six individual subjects, at $p < 0.001$.

In each of the individually-defined ROIs, we extracted the magnitude of the response to the two conditions in the biological motion experiment, and tested the difference between these conditions using repeated measures ANOVAs and t-tests.

Figure 2a shows the activation in left and right EBA and hMT+ for the two conditions in the biological motion experiment. A three-way repeated measures ANOVA (hemisphere x ROI x condition) revealed no significant interactions with hemisphere. ROI interacted significantly with condition ($F_{1,17} = 16.6, p < 0.001$), indicating a stronger effect of biological motion in EBA compared to hMT+. Paired-sample t-tests, however, showed that in each individual region the effect of biological motion was highly significant (left EBA: $t_{17} = 6.5, p < 0.001$, right EBA: $t_{17} = 5.3, p < 0.001$, left hMT+: $t_{17} = 4.2, p < 0.001$, right hMT+: $t_{17} = 5.3, p < 0.001$). Notably, a significant effect of biological motion was also found when the biological motion effect was measured from only the single most-selective (peak) voxel of each individual's ROIs (left EBA: $t_{17} = 7.3, p < 0.001$, right EBA: $t_{17} = 5.1, p < 0.001$, left hMT+: $t_{17} = 4.5, p < 0.001$, right hMT+: $t_{17} = 3.0, p < 0.01$).

Figure 2b shows the analogous results from right FBA and FFA. A two-way repeated measures ANOVA (ROI x condition) showed a significant interaction between ROI and condition ($F_{1,17} = 6.0, p < 0.05$), indicating a stronger effect of biological motion in FBA compared to FFA. Paired-sample t-tests showed that in both regions the effect of biological motion was significant (FBA: $t_{17} = 4.2, p < 0.001$, FFA: $t_{17} = 2.6, p < 0.05$). A significant effect of biological motion was also found when each ROI was defined by only the most selective voxel (FBA: $t_{17} = 3.4, p < 0.005$, FFA: $t_{17} = 3.2, p < 0.01$).

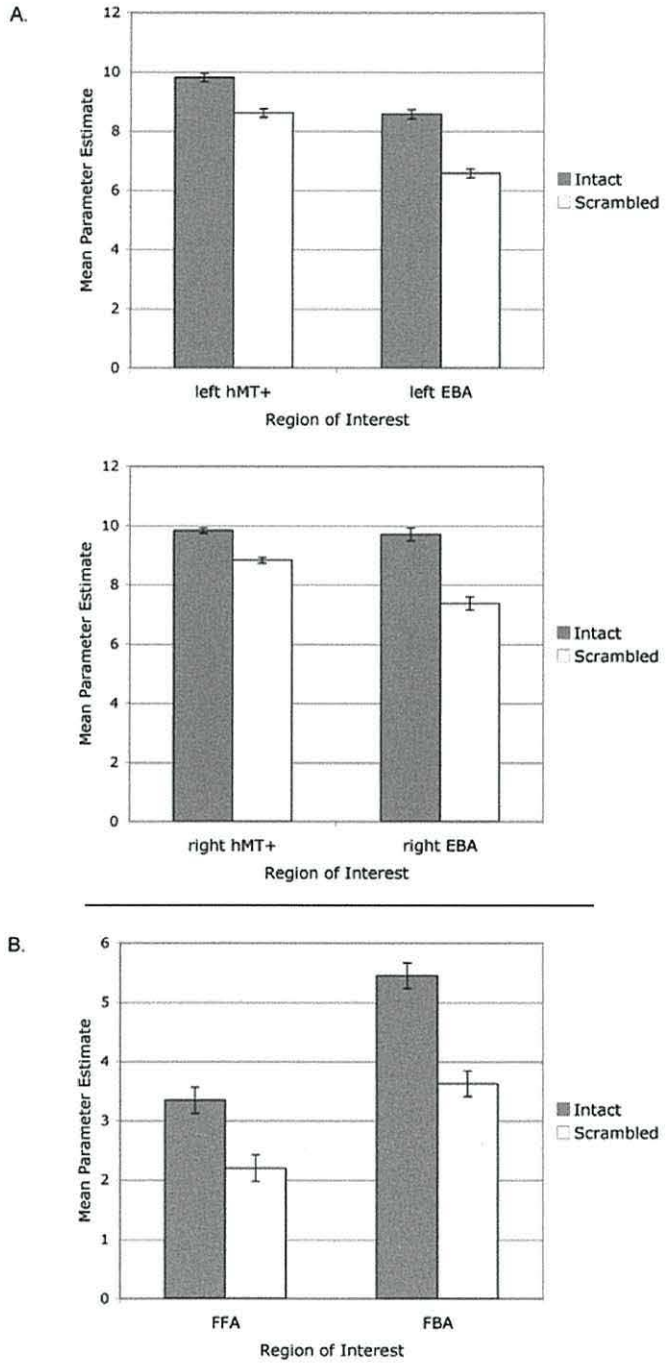


Figure 2. Biological motion activation in regions of interest

(A) Left hMT+ and left EBA (top panel), right hMT+ and right EBA (middle panel). (B) FBA and FFA (bottom panel). Error bars reflect within-subject s.e.m. (Loftus and Masson, 1994).

Interim summary

Thus left and right EBA, left and right hMT+, right FBA, and right FFA all show a strongly selective response to biological motion animations. This holds true whether the ROIs are defined as clusters of significant voxels surrounding (and including) the peak voxel, or as just the peak alone. One explanation of these results is that the body-selective, motion-selective, and face-selective neurons in these regions are all strongly activated by moving point light figures (versus control), albeit to a different extent. This explanation would require a functional account of how neurons generally selective for visual motion, and other neurons selective for faces, contribute to biological motion perception. An alternative explanation, with substantially different functional implications, is that the biological motion activation in hMT+ and FFA is entirely due to the existence of body-selectivity in these regions – perhaps in the form of body-selective neurons interspersed among face- or motion-selective neurons. On this account, motion selectivity and face selectivity are unrelated to biological motion selectivity.

Voxelwise correlations

To distinguish between these explanations, we determined the voxel-by-voxel correlations between biological-motion selectivity and selectivity for the localizer stimuli within each individually-defined ROI (see Figure 3a for an overview). For example, we reasoned that if motion-selective area hMT+ processes biological motion, voxels within hMT+ that are relatively strongly motion-selective (indicating that the voxel contains a relatively large number of motion-selective neurons) should also be relatively strongly

selective for biological motion. In contrast, voxels that show weaker motion selectivity should also be less selective for biological motion. In other words, we would expect a positive voxelwise correlation between motion selectivity and biological motion selectivity in hMT+ if 1) the two conditions activate the same neurons and 2) the variation in selectivity across voxels is stable and reflects variations in the proportions of neurons exhibiting different kinds of selectivity. (For a similar argument see Peelen and Downing, 2005c). Likewise, the same logic can be applied to the posterior fusiform gyrus activations. If the face-selective neurons of the FFA are involved in processing biological motion, there should be a positive voxelwise correlation in this region between face selectivity and biological motion selectivity. In contrast, if biological motion activation in the FFA is explained by the presence of body-selective neurons within the voxels assigned to that ROI, we would expect a positive correlation only between body selectivity and biological motion selectivity in this region.

Note that in the above discussion we assume that each ROI consists of body-, face- and motion-selective neurons, and then ask whether one of these populations is particularly engaged by biological motion. As we note in the Discussion, the same logic applies when we consider a less extreme scenario, in which a single population of neurons responds to varying degrees, and with different patterns across neurons, to the different stimulus types.

For each voxel in each ROI, on an individual subject basis, we calculated the biological motion selectivity (expressed by a t-value for each voxel). Then, we correlated these t-values with t-values reflecting the motion and body selectivity in hMT+ and EBA,

and body and face selectivity in FFA and FBA. The average correlations were then tested against zero, with subject as the random factor.

Voxel by voxel, biological motion selectivity was significantly correlated with body selectivity in all ROIs: left EBA ($r = 0.30$, $t_{17} = 5.4$, $p < 0.001$), right EBA ($r = 0.38$, $t_{17} = 5.8$, $p < 0.001$), left hMT+ ($r = 0.30$, $t_{17} = 3.6$, $p < 0.005$), right hMT+ ($r = 0.39$, $t_{17} = 5.0$, $p < 0.001$), FFA ($r = 0.31$, $t_{17} = 3.0$, $p < 0.01$), and FBA ($r = 0.14$, $t_{17} = 2.3$, $p < 0.05$). In contrast, biological motion selectivity did not correlate with motion selectivity in any of the pITS ROIs: left EBA ($r = -0.18$, $t_{17} = -2.0$, $p = 0.06$), right EBA ($r = 0.07$, $t_{17} = 0.7$, $p = 0.47$), left hMT+ ($r = 0.01$, $t_{17} = 0.1$, $p = 0.95$) and right hMT+ ($r = 0.00$, $t_{17} = 0.1$, $p = 0.92$). Nor was there a significant correlation with face selectivity in the pFG ROIs: FFA ($r = 0.13$, $t_{17} = 1.7$, $p = 0.10$), FBA ($r = 0.10$, $t_{17} = 0.8$, $p = 0.41$). Thus, in each region individually, biological motion selectivity was related to body selectivity but not motion or face selectivity.

To further generalize and verify these findings, we defined the union of EBA and hMT+ (which will be labelled pITS), and the union of FBA and FFA (which will be labelled pFG). This allows us to test for the general relationship between selectivities in these cortical "neighborhoods", without regard to whether a voxel is assigned to a particular labelled area. A further advantage of this analysis is that it controls for any differences in the main effect of biological motion that might exist between ROIs. That is, for example, it could be that the somewhat smaller biological motion effect in hMT+ (as compared to EBA) artificially suppresses the correlation between motion and biological motion.

Both left and right pITS were strongly body-, motion-, and biological motion selective (all ps < 0.001). Similarly, pFG was strongly body-, face-, and biological motion selective (all ps < 0.005). Within these larger ROIs, we again correlated the voxelwise patterns of biological motion selectivity with the patterns of body, motion, and face selectivity. Figure 3b gives the results of these analyses. Again, in all regions biological motion selectivity correlated positively with body selectivity (left pITS: $r = 0.47$, $t_{17} = 5.7$, $p < 0.001$; right pITS: $r = 0.57$, $t_{17} = 7.4$, $p < 0.001$; pFG: $r = 0.38$, $t_{17} = 4.2$, $p < 0.001$), but not (or negatively) with motion selectivity (left pITS: $r = -0.34$, $t_{17} = -4.8$, $p < 0.001$; right pITS: $r = -0.31$, $t_{17} = -3.9$, $p < 0.005$) and face selectivity ($r = 0.05$, $t_{17} = 0.2$, $p = 0.83$). See Figure 4, 5, and 6 for a graphical representation, in the form of scatterplots, of these results.

The negative correlations between motion selectivity and biological motion selectivity in pITS observed in the preceding analysis indicates that within the combined pITS ROI, EBA and hMT+ are to some degree discrete: voxels that are highly body- (and indeed biological motion-) selective tend to be non-responsive to motion, and vice versa. Note that this relationship only becomes apparent due to the use of the present pattern analysis: As described above, even when the single most body- or motion-selective peak voxel from each region is considered, both ROIs show highly significant selectivity for biological motion.

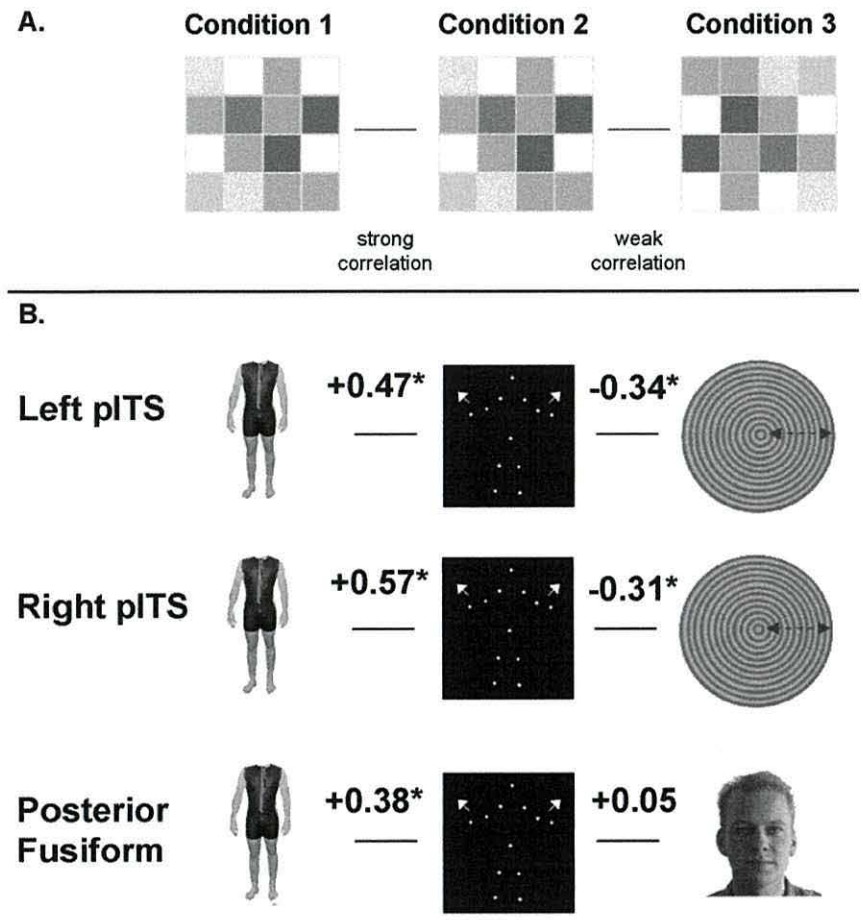
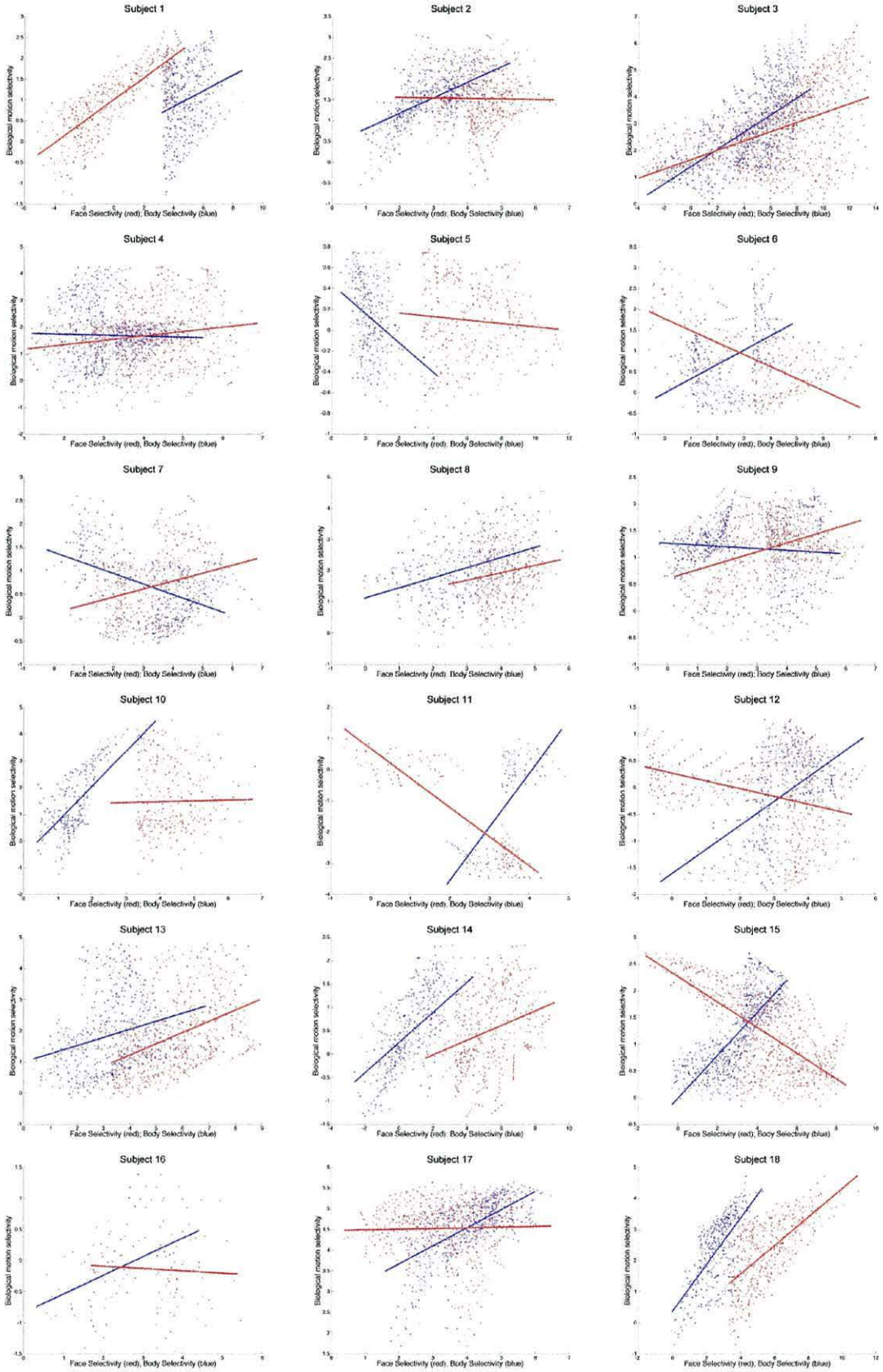
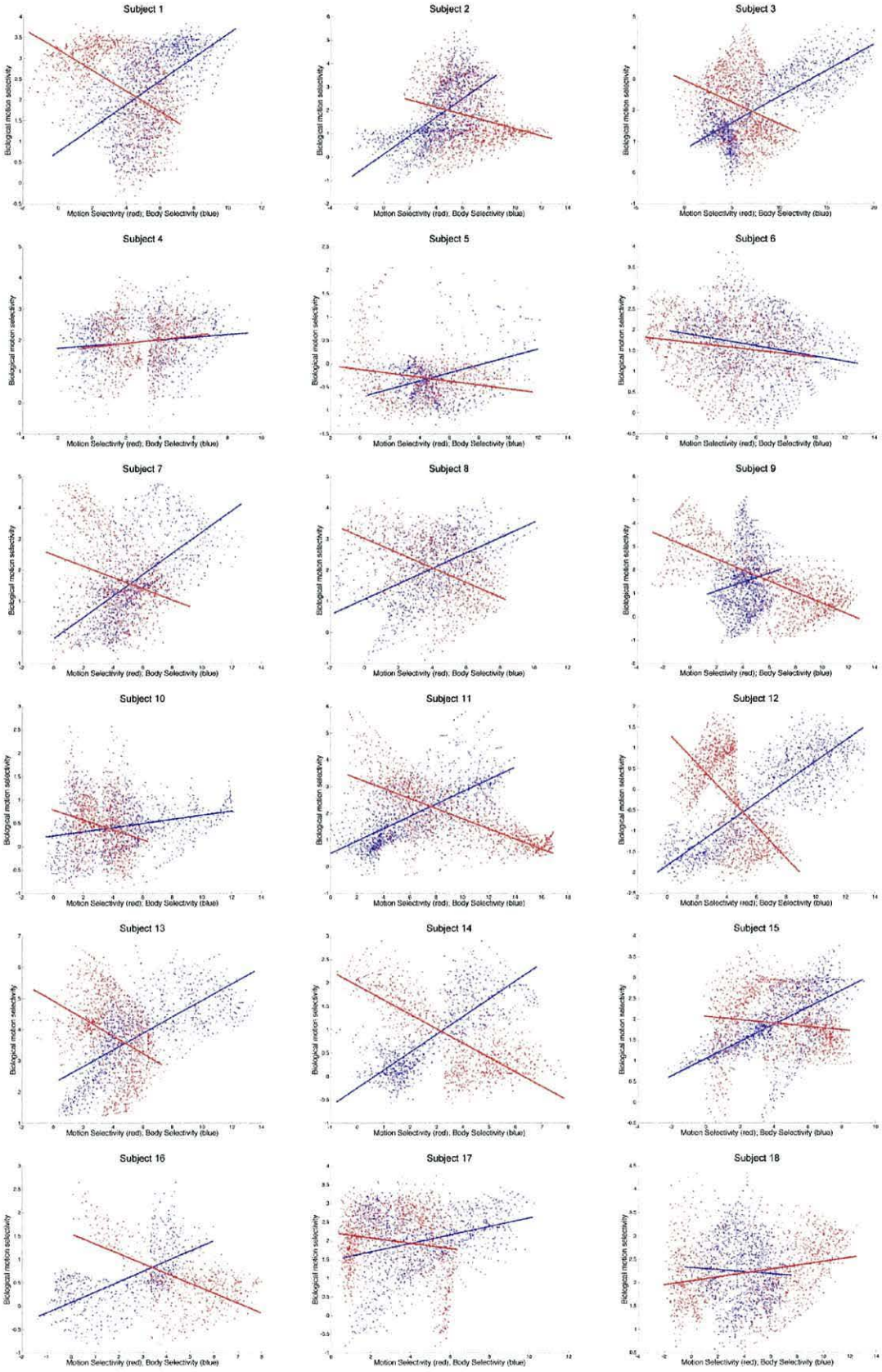


Figure 3. Voxelwise correlations

(A) Schematic overview of the voxelwise correlation method: activations to different conditions were correlated voxel-by-voxel. Conditions 1 and 2 elicit a similar activation pattern and are therefore highly correlated. Conditions 2 and 3 have a dissimilar activation pattern and are not (or negatively) correlated. (B) Voxelwise correlations between body, face and motion selectivity with biological motion selectivity in posterior ITS and posterior fusiform regions. Body selectivity correlated significantly with biological motion selectivity in all regions. Motion and face selectivity did not correlate, or correlated negatively, with biological motion selectivity.


Subject 1
Subject 2
Subject 3
Subject 4
Subject 5
Subject 6
Subject 7
Subject 8
Subject 9
Subject 10
Subject 11
Subject 12
Subject 13
Subject 14
Subject 15
Subject 16
Subject 17
Subject 18
Biological motion selectivity
Face Selectivity (red); Body Selectivity (blue)


Subject 1
Subject 2
Subject 3
Subject 4
Subject 5
Subject 6
Subject 7
Subject 8
Subject 9
Subject 10
Subject 11
Subject 12
Subject 13
Subject 14
Subject 15
Subject 16
Subject 17
Subject 18
Biological motion selectivity
Motion Selectivity (red); Body Selectivity (blue)

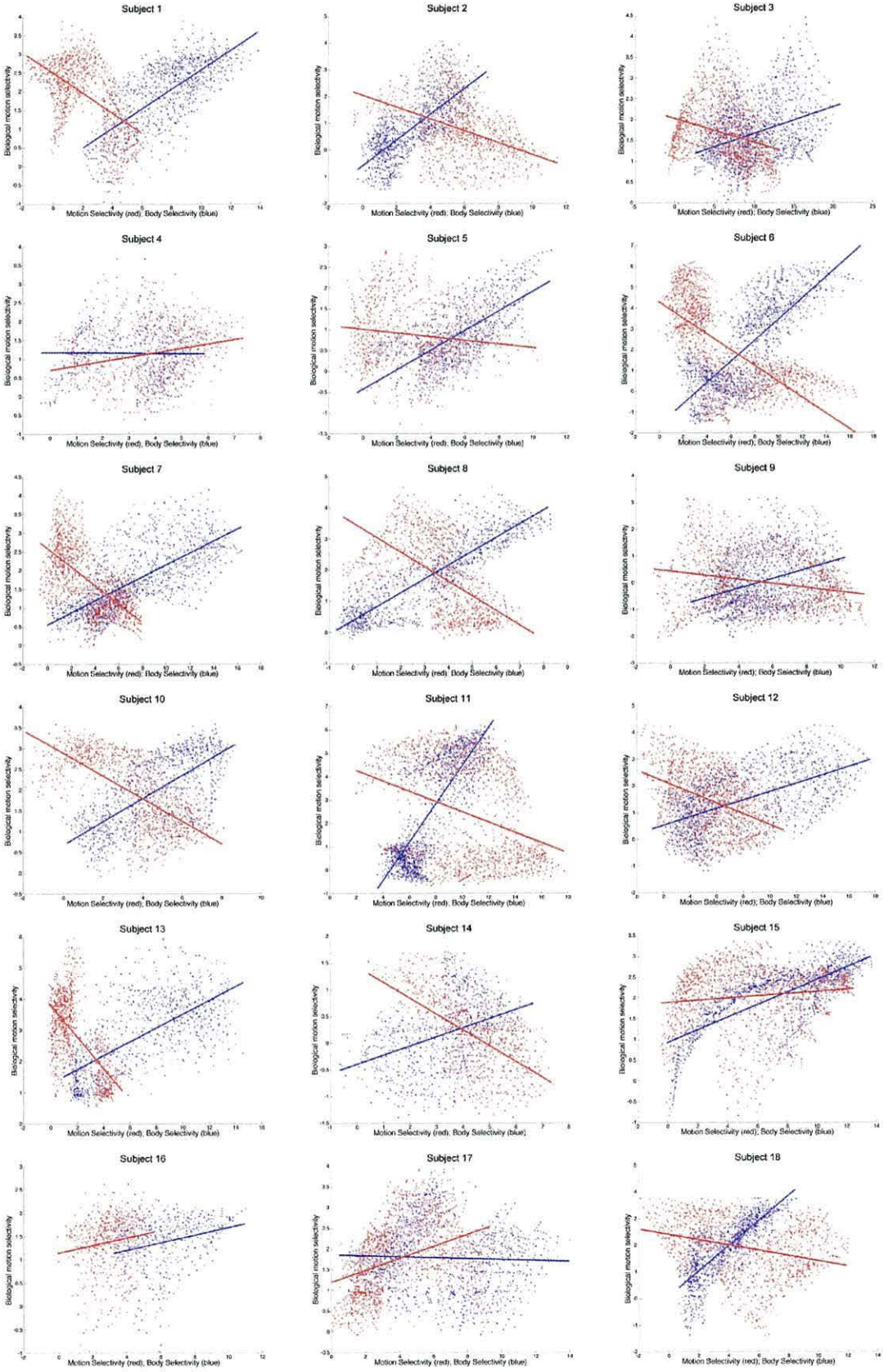

Subject 1
Subject 2
Subject 3
Subject 4
Subject 5
Subject 6
Subject 7
Subject 8
Subject 9
Subject 10
Subject 11
Subject 12
Subject 13
Subject 14
Subject 15
Subject 16
Subject 17
Subject 18
Biological motion selectivity
Motion Selectivity (red); Body Selectivity (blue)

Figure 4. Scatterplots, for each participant, showing the relationship between biological motion selectivity (y-axis) and both face-selectivity (x-axis; red dots) and body-selectivity (x-axis; blue dots). Voxels were extracted from individually-defined right hemisphere posterior fusiform regions of interest (ROIs). ROIs were composed of the union of voxels in the fusiform face area (FFA) and fusiform body area (FBA). Regression lines illustrate the linear relationship (if any) between the different types of selectivity. The correlations illustrated here are summarised in Figure 3b.

Figure 5. Scatterplots, for each participant, showing the relationship between biological motion selectivity (y-axis) and both motion-selectivity (x-axis; red dots) and body-selectivity (x-axis; blue dots). Voxels were extracted from individually-defined left hemisphere posterior inferior temporal sulcus regions of interest (ROIs). ROIs were composed of the union of voxels in the extrastriate body area (EBA) and human MT+. Regression lines illustrate the linear relationship (if any) between the different types of selectivity. The correlations illustrated here are summarised in Figure 3b.

Figure 6. As in Figure 5, except for right hemisphere pITS region.

Finally, to assess the independent predictive value of the voxelwise patterns of motion, body, and face selectivity to the patterns of biological motion selectivity in the three regions (left and right pITS, and pFG), we performed multiple regression analyses. For each subject and region, a regression model was tested with biological motion selectivity as the dependent variable, and body and motion selectivity (pITS), or body and face selectivity (pFG) as predictors. The obtained betas were accumulated across subjects and tested against zero. This analysis tests, for each area, whether body selectivity is a significant predictor of biological motion selectivity while simultaneously taking into account any shared variance with face or motion selectivity. Biological motion selectivity

could be predicted by body selectivity in all regions (left pITS: $\beta = 0.45$, $t_{17} = 6.1$, $p < 0.001$; right pITS: $\beta = 0.54$, $t_{17} = 9.8$, $p < 0.001$; pFG: $\beta = 0.31$, $t_{17} = 3.4$, $p < 0.005$) but not by motion selectivity (left pITS: $\beta = -0.17$, $t_{17} = 2.0$, $p = 0.06$; right pITS: $\beta = -0.11$, $t_{17} = 1.4$, $p = 0.18$) or by face selectivity (pFG: $\beta = 0.13$, $t_{17} = 1.3$, $p = 0.22$).

DISCUSSION

Our findings support three main conclusions, each of which we discuss in turn below. First, we resolve a longstanding ambiguity in studies of the neural basis of biological motion perception, by identifying and interpreting the source of lateral ITS and posterior FG biological motion activations. Thus this work significantly clarifies our emerging picture of how the human brain makes sense of the appearance and actions of other individuals. Second, we provide new evidence for a highly-selective, focal representation of the human body that closely overlaps, but is functionally separate from, the fusiform face area. This finding has important implications for interpreting the functional organization of this region. Finally, our results illustrate the power of combining a functional ROI approach with analyses of the response patterns within different ROIs. By localizing multiple, adjacent areas of interest and relating various types of selectivity on a voxel-by-voxel basis, we were able to draw conclusions that would be impossible with typical whole-brain group-average or functional ROI analyses.

Biological motion perception

Numerous previous studies have supported the notion of a network of areas involved in the perception of other individuals. With respect to biological motion, the focus has been almost entirely on pSTS, to the exclusion of other posterior regions engaged by the same stimuli. Our results show that visual areas involved in analyzing the form of the human body (EBA, FBA) are selectively activated by sparse movement patterns that induce the percept of a person performing an action. In these regions, there was not only a global preference for point-light actors compared to scrambled controls, but also a strong voxel-by-voxel correlation between body selectivity and biological-motion selectivity. Other areas (hMT+ and FFA) that substantially overlapped these regions also showed a significant selective activation to the biological motion displays. Strikingly, however, the voxelwise variation in response to these regions' preferred stimuli (simple motion and faces) bore no relationship to biological motion selectivity, while variation in body selectivity did. From these results we conclude that, in spite of the apparent biological motion selectivity of hMT+ and FFA, these regions play no functional role in biological motion perception. Instead we attribute the biological motion response in these two general cortical areas to the EBA and FBA, respectively.

What do our results suggest about the pITS and pFG regions at the level of individual neurons? To use pITS as an example, one possibility is that the region contains two kinds of neurons, which are highly selective either for visual features of the human body, or for visual motion. Alternatively, the distinction between motion- and body-selective neurons may be one more of degree than of kind. On this account, pITS neurons would respond to varying degrees to both visual motion and to visual aspects of the

human body. Our results indicate that if this were the case, the distribution of these two kinds of selectivity would be largely independent. That is, variations in body selectivity from neuron to neuron would be unrelated to (and not predictable by) variations in motion selectivity -- indicating functional independence of motion and body selectivity, even if these two functions are not divided in an absolute sense between two neural populations.

On either account, body- and motion-selectivity would have to be interleaved on a fairly fine scale (relative to the resolution of fMRI), to explain our findings of biological motion selectivity in hMT+ as well as EBA ROIs. The results of our pattern analyses, however, suggest consistent variation across voxels in the relative proportion of neural selectivity for bodies and for visual motion (e.g. Haynes and Rees, 2005; Kamitani and Tong, 2005).

What is the functional difference between the various posterior areas engaged by biological motion? Importantly, pSTS is not strongly selective for static bodies (e.g. Grossman et al., 2002), in contrast to the EBA and FBA. This suggests that pSTS is activated specifically by particular motion patterns, instead of the form that they imply. This is consistent with a study by Beauchamp et al. (2002), who tested the neural response to the same bodies moving in either an articulated or unarticulated fashion. Significantly more activation was found in pSTS during articulated body movements, suggesting that the type of movement is a key predictor for pSTS activation. In contrast, areas EBA and FBA have been shown to selectively respond to static bodies even when they are depicted by very minimal visual cues, for instance a few lines forming a "stick figure" (Downing, 2001; Peelen and Downing, 2005a). In other words, the presence of

the form of the body seems crucial for these areas, much more so than the presence of specific low-level visual features.

Taking these results together, we speculate that the EBA and FBA are largely "naïve" about the patterns of changing posture that comprise biological actions, but instead simply respond to the presence of the form of the body. The pSTS, in contrast, integrates information over time, and activates in response to movements that are biologically plausible. These suggestions are generally consistent with a recent computational model of biological motion perception, which proposes that pSTS integrates form information from a "ventral" pathway (which includes EBA) and motion information from a "dorsal" pathway (Giese and Poggio, 2003). Note, however, that this model does not take into account the recently-discovered FBA; we turn next to a closer discussion of this region.

Body selectivity and the posterior fusiform gyrus

Body-selective responses in FBA, but not face-selective responses in overlapping area FFA, showed a strong relation to selectivity for biological motion. This finding is important because it shows that areas that are anatomically overlapping can be functionally dissociated. A recent study, using high resolution (1.4 x 1.4 x 2.0mm) fMRI, found that pFG body and face responses could also be partly distinguished spatially, in that in many subjects voxels could be identified that showed either body or face selectivity, but not both (Schwarzlose et al., 2005). Notably, areas of overlap remained between face- and body-selective regions. The present results suggest that body and face

selectivity may be independent within the overlapping region, and offers a method to test this possibility with high-resolution data.

The face selectivity of the posterior fusiform gyrus has been the matter of much recent debate (Kanwisher, 2000; Tarr and Gauthier, 2000). On one account, the FFA is not selective for faces as such, but is instead involved in a cognitive process that is usually most strongly employed by faces (Tarr and Gauthier, 2000). This cognitive process has been qualitatively described as sub-ordinate discrimination of highly similar objects for which one has substantial expertise. This account of the FFA could potentially explain the strong body selectivity in this general region (Peelen and Downing, 2005a, Schwarzlose et al., 2005), as the perceptual processes involving bodies are in some ways similar to those involving faces (this is illustrated, for example, by the inversion effect found for both bodies and faces [Reed et al., 2003]). Importantly, the present evidence that body- and face-selective fusiform regions can be functionally dissociated refines our view of the properties of this region. We propose that it contains (at least) two functionally distinct domain-specific representations, overlapping at a relatively fine scale. These populations may share in common the property that they are suited to discriminating highly similar objects, which would be consistent with a domain-general account of the region as a whole.

Analysis of multiple overlapping regions of interest

As we have shown for two different broad brain regions (posterior inferior temporal sulcus and posterior fusiform gyrus), group-average coordinates of functionally quite different regions can be very close together. As a result, great caution should be

taken in functionally labeling an activated region (e.g. "FFA") based on group-average coordinates, especially when the coordinates of this area come from a different group of subjects or a different study altogether.

One approach that researchers have taken to overcome this problem is to functionally localize ROIs in each subject, thus avoiding the problems that arise from inter-subject averaging. However, our results show that this may not be sufficient to avoid false conclusions. Indeed, based simply on the overall amplitude differences that we found between intact and scrambled biological motion sequences (which were highly significant in each ROI individually, even when only peak voxels were considered), we might have falsely concluded that body-selective, motion-selective and face-selective neurons are all involved in biological motion perception.

This highlights the importance of localizing not only regions of interest critical for one's hypothesis, but also other known nearby regions. When these regions overlap, as in the case of hMT+/EBA and FFA/FBA, this localization must be done within subjects, and ideally within sessions, to maximize the ability to distinguish functional regions. Finally, and critically, the pattern of responses to a contrast of interest, across the voxels that comprise each ROI, carries valuable information (Cox and Savoy, 2003; Haxby et al., 2001; Haynes and Rees, 2005; Kamitani and Tong, 2005). In the present study, analyzing these patterns allowed us to disentangle the sources of biological motion selectivity as observed at the aggregate ROI level. Because of the unique voxelwise relationship between body selectivity and biological motion selectivity, we can attribute the biological motion effects in hMT+ and FFA to the presence of body-selectivity in the voxels of those ROIs.

Conclusion

To summarize, we have used a combination of multiple within-subject ROI definition and voxelwise pattern analyses to elucidate the functional network of brain regions involved in the analysis of biological motion. We believe this approach will prove useful both in further studies on the neural basis of "social vision", and more generally in studies of other regions where multiple functional areas occupy overlapping cortical territory.

Chapter 5:

The role of the extrastriate body area in action perception

ABSTRACT

Numerous cortical regions respond to aspects of the human form and its actions. What is the contribution of the extrastriate body area (EBA) to this network? In particular, is the EBA involved in constructing a dynamic representation of observed actions? We scanned 16 participants with fMRI while they viewed two kinds of stimulus sequences. In the *coherent* condition, static frames from a movie of a single, intransitive whole-body action were presented in the correct order. In the *incoherent* condition, a series of frames from multiple actions (involving one actor) were presented. ROI analyses showed that the EBA, unlike area MT+ and the posterior superior temporal sulcus, responded more to the incoherent than to the coherent conditions. Whole brain analyses revealed increased activation to the coherent sequences in parietal and frontal regions that have been implicated in the observation and control of movement. We suggest that the EBA response adapts when succeeding images depict relatively similar postures (coherent condition) compared to relatively different postures (incoherent condition). We propose that the EBA plays a unique role in the perception of action, by representing the static structure, rather than dynamic aspects, of the human form.

INTRODUCTION

One of the most important functions of vision is to provide information about conspecifics – that is, to inform us about the identity, actions, and mental states of other people around us. Extensive research has focused on how the brain extracts this information from the visual input. Converging evidence has accumulated for face-specific brain mechanisms in humans and other primates (Behrmann and Avidan, 2005; Carmel and Bentin, 2002; Desimone, Albright, Gross, and Bruce, 1984; Haxby, Hoffman, and Gobbini, 2002; Kanwisher, 2000). In parallel, other studies have examined the visual analysis of the form and movement of others' bodies. A key observation from this work, which is reviewed briefly below, is that multiple areas of the human brain respond selectively to some aspect of the human form and its actions.

In the monkey, neurons in the superior temporal sulcus, particularly the anterior portion, are driven by head and body movements (e.g. Jellema and Perrett, 2003; Perrett et al., 1985). In humans, regions of the right posterior STS (pSTS) respond selectively in fMRI to human biological motions, whether depicted in movies (Calvo-Merino, Glaser, Grèzes, Passingham, and Haggard, 2005), animations (Pelphrey, Mitchell, McKeown, Goldstein, Allison, and McCarthy, 2003), or versions of Johansson's (1973) schematic "point light walker" displays (Grèzes, Fonlupt, Bertenthal, Delon-Martin, Segebarth, and Decety, 2001; Grossman, Donnelly, Price, Pickens, Morgan, Neighbor, and Blake, 2000; see also Allison, Puce, and McCarthy, 2000 for review). The extrastriate body area (EBA; Downing, Jiang, Shuman, and Kanwisher, 2001), located in the posterior inferior temporal sulcus, responds generally and selectively to images of human bodies and body

parts (but no more to faces than to objects). Notably this region also responds significantly more to point-light animations than to scrambled controls (Downing et al., 2001b; see also Michels, Lappe, and Vaina, 2005). Finally, a region of the posterior fusiform gyrus responds strongly and selectively to images of the human body without the face (fusiform body area or FBA; Peelen and Downing, 2005a; Schwarzlose, Baker, and Kanwisher, 2005). In this region, the group average activation to bodies (vs. objects) is nearly as strong as that to faces. However, on grounds of a functional dissociation (Peelen and Downing, 2005a) and high-resolution functional imaging (Schwarzlose et al., 2005) a case has been made for distinct face and body representations in fusiform gyrus. A strong response to point-light displays has also been observed in a similar region (Grossman, Blake, and Kim, 2004; Santi, Servos, Vatikiotis-Bateson, Kuratate, and Munhall, 2003), consistent with activation of the FBA by these schematic stimuli. Finally, the posterior fusiform gyrus shows enhanced activation to emotional, relative to neutral, body postures (Hadjikhani and de Gelder, 2003). This may reflect modulation of FBA activity by attentional or motivational factors.

Action observation, particularly of actions involving interaction with an object, also produces enhanced single-unit activity in the parietal and inferior frontal cortices of monkeys (di Pellegrino, Fadiga, Fogassi, Gallese, and Rizzolatti, 1992; Fogassi, Ferrari, Gesierich, Rozzi, Chersi, and Rizzolatti, 2005), and enhanced BOLD activity in similar regions of the human brain (Buccino, Vogt, Ritzl, Fink, Zilles, Freund, and Rizzolatti, 2004; Iacoboni, Woods, Brass, Bekkering, Mazziotta, and Rizzolatti, 1999). In the monkey, these neurons respond both when the animal makes a particular movement, and when the monkey observes a similar movement. For this reason, these neurons have been

termed "mirror neurons" and have been implicated in imitation learning and in the representation of others' intentions (see Rizzolatti and Craighero, 2004, for a review).

Thus a number of relatively distant cortical areas all respond selectively to some aspect of the visual appearance or actions of other people. An important aim for research on action perception is to distinguish among the functional properties of these regions. The present study moves towards this goal, by attempting to determine how the EBA contributes to this network of areas. Specifically, we ask whether the EBA is directly involved in representing the dynamic aspects of human actions – that is, in the integration of changes in the configuration of a person's body over time. This might be expected if the cortical areas reviewed above interact in a mutually-reinforcing manner, in which the EBA's activity is modulated by feedback from the fronto-parietal motor control areas. An account of this sort has been offered (Jeannerod, 2004) to explain the finding that activity in the EBA increases during the (blind) performance of simple visually-guided motor movements (Astafiev et al., 2004; Jackson et al., 2006; but see Peelen and Downing, 2005c for an alternative interpretation). This account could be extended to propose that the EBA receives efferent information about not only executed actions but also (via "mirror" systems) observed actions. Thus, on what we call the "dynamic representation" hypothesis, we would expect the response of the EBA (relative to appropriate controls) to mirror that of the parietal, frontal, and superior temporal regions that respond strongly to (and indeed produce) biological actions.

An alternative – the "static representation" hypothesis -- is that the EBA, while playing a role in the perception of others' bodies, is not directly involved in the dynamic representation of biological actions. Instead, on this account the EBA responds simply to

the visual presence of the human form, and is "naive" to higher-level manipulations such as the dynamic context in which the body is viewed (see also Chan et al., 2004 for a similar hypothesis). Consistent with the "static representation" hypothesis is a recent model of biological motion perception (Giese and Poggio, 2003). This model, based on computational considerations, neuroimaging studies, and neurophysiological findings, proposes that the EBA's role is one of representing static "snapshots" of the individual postures that comprise whole-body actions.

How can we distinguish between the static and dynamic representation hypotheses? Much of the neuroimaging research on biological action perception has used point-light animations of simple actions, compared to controls, as experimental stimuli. These stimuli are valuable in that they can test the response of brain regions to biological motion *per se*, in the absence of other visual cues. However, a limitation of this approach for studying the EBA for the present purposes is that in the experimental condition, one has the percept of a human figure as a result of structure-from-motion. This percept is, by definition, not present in the control conditions typically used. (These include both simple linear or radial motion patterns, and "scrambled" point-light animations, in which individual point movements are retained but the starting points determined randomly). Thus the differential response to these stimuli in the EBA is ambiguous – perhaps the EBA processes biological motion, but perhaps it simply responds more to the experimental condition because it elicits the percept of a human body. Indeed, previous work has shown that highly schematic depictions of the human body (e.g. silhouettes or stick figures) can selectively activate the EBA (Downing et al., 2001b; Peelen and Downing, 2005a).

For this reason, we developed a test of action perception in which the presence of a human figure is constant in both critical conditions. We scanned participants while they viewed two types of displays containing human actors (see Figure 1). The raw materials for both conditions were still frames taken from short movies depicting simple intransitive actions. In the *coherent* condition, the frames were presented in the correct order. In the *incoherent* condition, frames from different movies of one actor were combined to make a single meaningless sequence. In both conditions the displays were presented at a sufficiently slow rate to disrupt the percept of visual motion. Likewise, at this slow rate of presentation the transitions between frames did not give rise to the percept of biomechanically-impossible movements (cf. Shiffrar and Freyd, 1990). If the "dynamic representation" account of the EBA is correct, we should expect a greater response to the coherent sequences than to the incoherent sequences, in spite of equivalence in the two conditions at the level of individual frames. In contrast, if the "static representation" hypothesis is correct, we would expect no such difference.

We analyzed the data in two ways. First, we used a functional localiser approach in order to identify the right hemisphere EBA in each participant. For comparison, the right pSTS, FBA, and motion-selective MT+ were also identified. We measured the neural response of each of these regions to the coherent and incoherent sequences. Second, in order to verify that our coherent sequences activated other brain regions that respond to human actions (e.g. in parietal and frontal cortex), we performed a whole-brain, group-average contrast of coherent and incoherent action sequences.

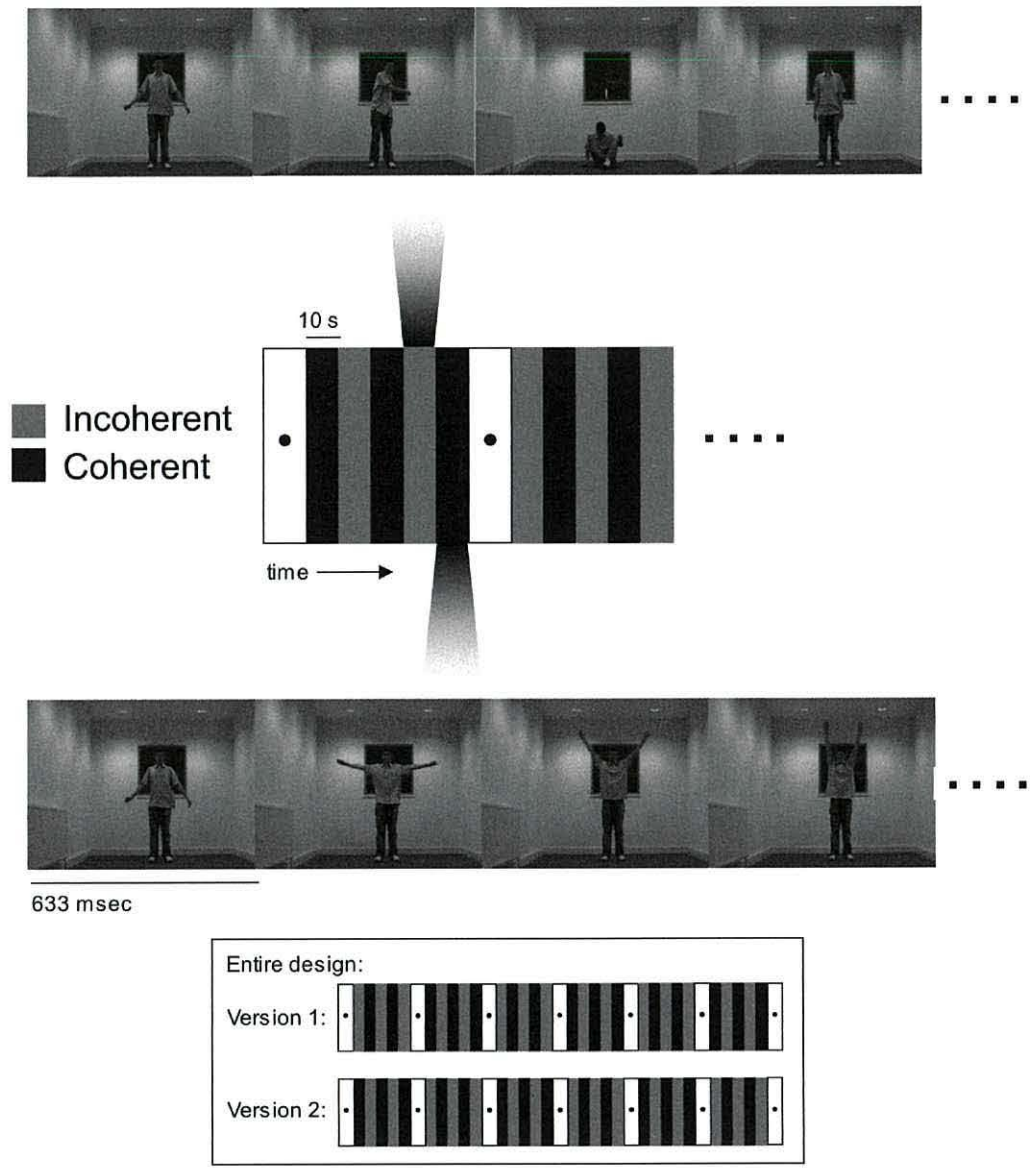


Figure 1. Illustration of the experimental paradigm, with details of the timecourse of the design and example stimuli.

MATERIALS AND METHODS

Subjects

Sixteen healthy adult volunteers (mean age = 23, range = 18-34; 9 female) were recruited from the University of Wales, Bangor community. Participants satisfied all requirements in volunteer screening and gave informed consent approved by the School of Psychology at the University of Wales, Bangor and the North-West Wales Health Trust. Participation was compensated at £20 per session.

Design and Procedure

Localiser scans.

Each participant was scanned on a series of blocked-design localiser experiments, in order to identify *a priori* functional regions of interest with respect to individual brain anatomy. One localiser consisted of blocks of images of human faces, human bodies without heads, outdoor scenes, and handheld tools. Each condition was presented in four 15 sec blocks within one scan. In each block 20 different images from one category were presented (300 msec on / 450 msec off). Each participant was tested with two different versions of the localiser, counterbalancing for the order of the blocks. Participants performed a "1-back" repetition detection task during this localiser. Further details can be found elsewhere (Peelen and Downing, 2005b).

The localiser for area MT+ consisted of low-contrast, concentric rings that either slowly oscillated inwards and outwards, or, in separate blocks, remained static (Tootell, Reppas, Kwong, Malach, Born, Brady, Rosen, and Belliveau, 1995). Each condition was

presented for eight 15 sec blocks. Blocks alternated between moving and static, with a fixation-only block at the beginning and end of the scan, and after every fourth block. The localiser for posterior STS followed a similar design, with 16 sec blocks in which either point-light-walker animations or phased-scrambled versions of the same animations were presented (Grossman et al., 2000). Each animation lasted 667 msec, with a 333 msec blank interval between each stimulus.

Main Experiment.

Each of two actors (1 male, 1 female) were videotaped against a neutral background while performing 15 short actions (e.g. punching, kicking, bending, etc.; see Figure 1 for examples). The two actors performed different actions. The actors' whole bodies were always visible. None of the actions involved interacting with real objects, although a few of the actions implied interaction with an object (e.g. running a comb through the hair). After the movies were digitized, fifteen individual frames were taken from each action, with each frame separated from the next by approximately 200 msec. These images were used to build the sequences presented in the main experiment.

The design of the main experiment is illustrated in Figure 1. Each participant was scanned twice, with each scan comprising images of a different actor. There were six major blocks per scan, interrupted by blocks in which only a fixation point was visible for 15 sec. Each scan also began and ended with a 15 sec fixation block. Each of the six major blocks was divided into 5 minor blocks. In each minor block, coherent and incoherent 15-image sequences were presented in alternation. Half of the major blocks began with a coherent sequence, and half with an incoherent sequence. The initial condition of the run was counterbalanced across scans and participants.

Each image was presented for 633 msec, with no interval between images. The final image of each sequence was followed by a blank screen for 500 msec, resulting in a duration of 10 sec per sequence. In the coherent blocks, the 15 frames from a single action were presented in order. In the incoherent blocks, frames from different actions (but always depicting the same actor) were presented. The order of these frames was random, with the constraints that each frame appeared only once per scan, and that no frame was preceded or followed by a frame taken from the same original action movie. Each frame from each action (225 total) appeared exactly once in the coherent condition and once in the incoherent condition in a given scan. Thus the stimuli were matched between the coherent and incoherent conditions at the level of individual images. Participants were told to attend to the sequences but were otherwise given no explicit task.

Data Acquisition

A 1.5T Philips MRI scanner with a SENSE parallel head coil was used. For functional imaging, a single-shot EPI sequence was used (T2* weighted, gradient echo sequence, TE = 50 ms, flip angle 90°). Localisers were scanned with TR=3000 ms, 20-22 off-axial slices with no gap, voxel dimensions of 3.75x3.75x5 mm (N=10) or 3x3x4 mm (N=6). The main experiment was scanned with TR=3000 ms, 22 slices with no gap, 3.75x3.75x5 mm voxel dimensions, except for one participant (20 slices at 3x3x4 mm).

Data Analysis

Pre-processing and statistical analysis of MRI data was performed using BrainVoyager 4.9 (Brain Innovation, Maastricht, The Netherlands). Three dummy volumes were acquired before each scan in order to reduce possible effects of T1 saturation. Functional data were motion-corrected, low-frequency drifts were removed with a temporal high-pass filter (0.006 Hz), and spatial smoothing was applied with a 6mm FWHM filter. Functional data were manually co-registered with 3D anatomical T1 scans (1 x 1 x 1.3 mm resolution, resampled to 1x1x1 mm voxels with trilinear interpolation). The 3D anatomical scans were transformed into Talairach space (Talairach and Tournoux, 1988), and the parameters for this transformation were subsequently applied to the co-registered functional data, which were resampled to 1x1x1 mm voxels.

For each participant, general linear models were created for each localiser experiment and for the main experiment. One "boxcar" predictor (convolved with a standard model of the HRF) modelled each condition. Regressors were also included to account for differences in global signal across scans.

In each participant, the localiser scans were used to define the EBA by contrasting the response to human bodies with that to the average of faces, tools, and scenes. The FBA was defined by contrasting bodies against tools. We identified area MT+ by contrasting the response to moving concentric rings with that to static rings. Finally, the posterior STS ROI was identified by comparing the response to point-light walker stimuli with phase-scrambled versions of the same stimuli. Analyses were restricted to right hemisphere ROIs, on the basis of evidence that the EBA, FBA, and pSTS biological

motion regions are weaker or non-existent in the left hemisphere (Downing et al., 2001b; Peelen and Downing, 2005a,b; Allison et al., 2000).

For each ROI in each subject, the most significantly activated voxel was identified within a restricted part of cortex based on previously-reported anatomical locations (EBA: Peelen and Downing, 2005b; FBA: Peelen and Downing, 2005a; MT+: Dumoulin, Bittar, Kabani, Baker, Le Goualher, Bruce Pike, and Evans, 2000; pSTS: Grossman et al., 2000). ROIs were defined as the set of contiguous voxels that were significantly activated (all $p < 0.0001$ uncorrected, except pSTS: $p < 0.005$) within a 9 mm cube surrounding the peak voxel. This procedure was adopted for four reasons: 1) to ensure that regions were defined objectively; 2) to ensure that they were segregated from nearby selective activations; 3) to roughly equate the number of voxels included across different regions of interest; and 4) to ensure that only the most selective voxels were included in the ROI. Within each ROI in each subject, a further general linear model was then applied, modelling the response of the voxels in the region, in aggregate, to the coherent and incoherent conditions of the primary experiment. The regression weights from this GLM provided the basis for the ROI results described below.

Finally, a whole-brain, group average analysis was conducted on data from the main experiment. Fixed-effects contrasts were performed, at an uncorrected threshold of $p < 0.000005$, to test for regions more active in either the coherent or incoherent conditions. Only clusters > 100 mm^3 are reported for this analysis.

RESULTS

ROI analyses.

We successfully identified the right hemisphere ROIs in all participants, with the exception of pSTS, which was identified in 15/16 participants. The mean (with SEM) Talairach coordinates of the ROI peaks were as follows: EBA: [48 (1.3), -70 (1.5), 0 (1.5)]; pSTS: [54 (1.6), -46 (1.9), 12 (1.8)]; MT+: [44 (1.1), -67 (1.7), -3 (1.8)]; FBA: [41 (0.7), -43 (2.1), -19 (1.4)]. Further ROI analyses were conducted on the 15 participants who showed all ROIs.

Average parameter estimates for each ROI are given in Figure 2. An initial 2 (coherent or incoherent) X 4 (EBA, pSTS, MT+, or FBA) within-subjects ANOVA showed a significant main effect of ROI, $F(3,42) = 7.4$, $p < 0.001$, and, more important, a significant interaction of ROI with coherence, $F(3,42) = 6.2$, $p < 0.005$. The main comparisons of interest for the ROI analyses concerned whether the EBA showed similar or different effects of coherence as compared to pSTS, MT+, and FBA. Therefore we performed a 2 x 2 within-subjects ANOVA comparing the effects of coherence in the EBA and each of these regions. The comparison of the EBA and pSTS revealed a significant interaction, $F(1,14) = 15.7$, $p < 0.001$, and a main effect of ROI, $F(1,14) = 11.7$, $p < 0.005$. Follow-up t-tests showed that in the EBA, there was a significantly greater response to incoherent than to coherent sequences, $t(14) = 3.0$, $p < 0.01$. In pSTS, in contrast, the response was marginally greater to coherent than to incoherent sequences, $t(14) = 1.8$, $p < 0.1$. Area MT+ and the EBA also showed significantly different effects of coherence, $F(1,14) = 12.6$, $p < 0.005$. A simple effects test revealed no effect of

coherence in MT+, $t(14) = 0.03$, ns. Finally, the EBA and FBA showed a significant main effect of coherence, $F(1,14) = 6.2$, $p < 0.05$, but no interaction, $F(1,14) = 1.3$, $p = 0.28$. A simple effects test in the FBA alone showed no significant effect of coherence, $t(14) = 1.1$, ns. (The response of the FBA interacted significantly with pSTS, $F(1,14) = 8.5$, $p < 0.025$, but not with MT+, $F(1,14) = 1.1$, $p=0.30$.) In sum, in the EBA, incoherent sequences produced a larger response than coherent sequences. This pattern was reliably different from that observed in pSTS and MT+, but not FBA.

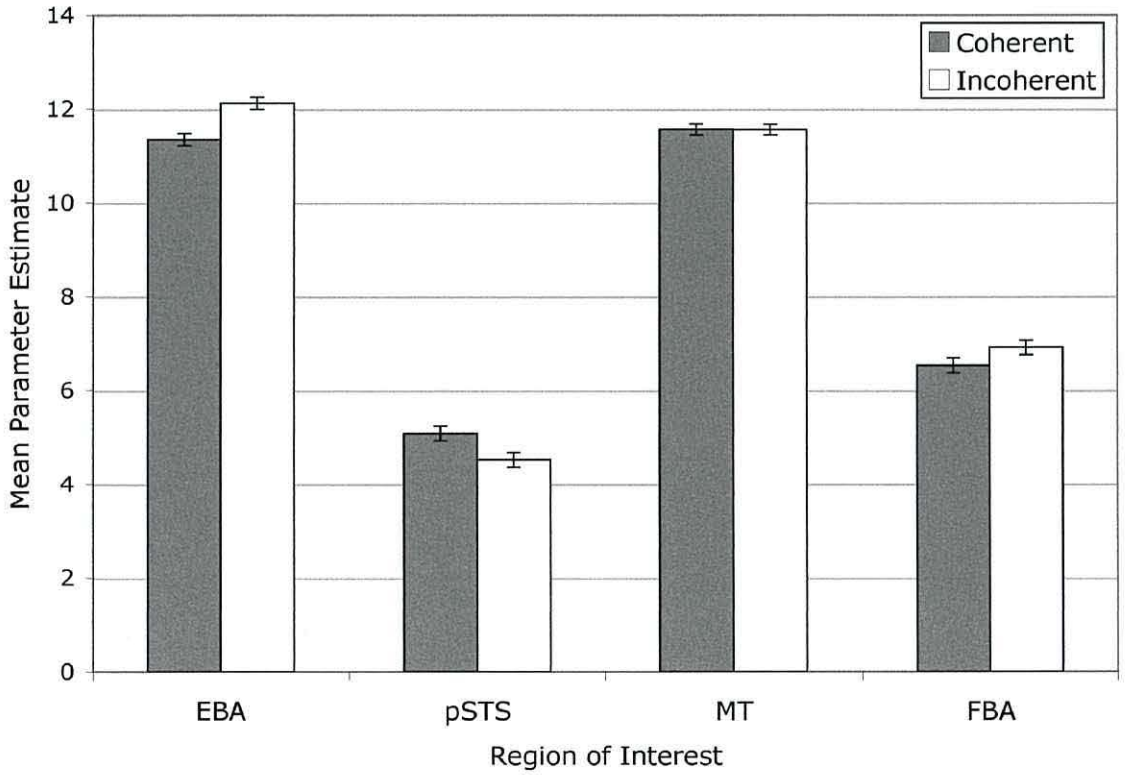


Figure 2. Results of the region of interest analysis. Mean parameter estimates are given for the response to coherent and incoherent sequences, in the right hemisphere extrastriate body area, posterior superior temporal sulcus, area MT+, and fusiform body area. Error bars reflect within-subjects standard error of the mean calculated for each ROI separately.

Whole-brain analyses.

The whole brain analysis was conducted on the data from all 16 participants. We observed significant increases to the coherent condition, relative to the incoherent condition, in the right parietal cortex, right ventral inferior frontal gyrus, and left occipitotemporal cortex (see Figure 3). A list of all activations greater than 100 mm^3 in extent is given in Table 1.

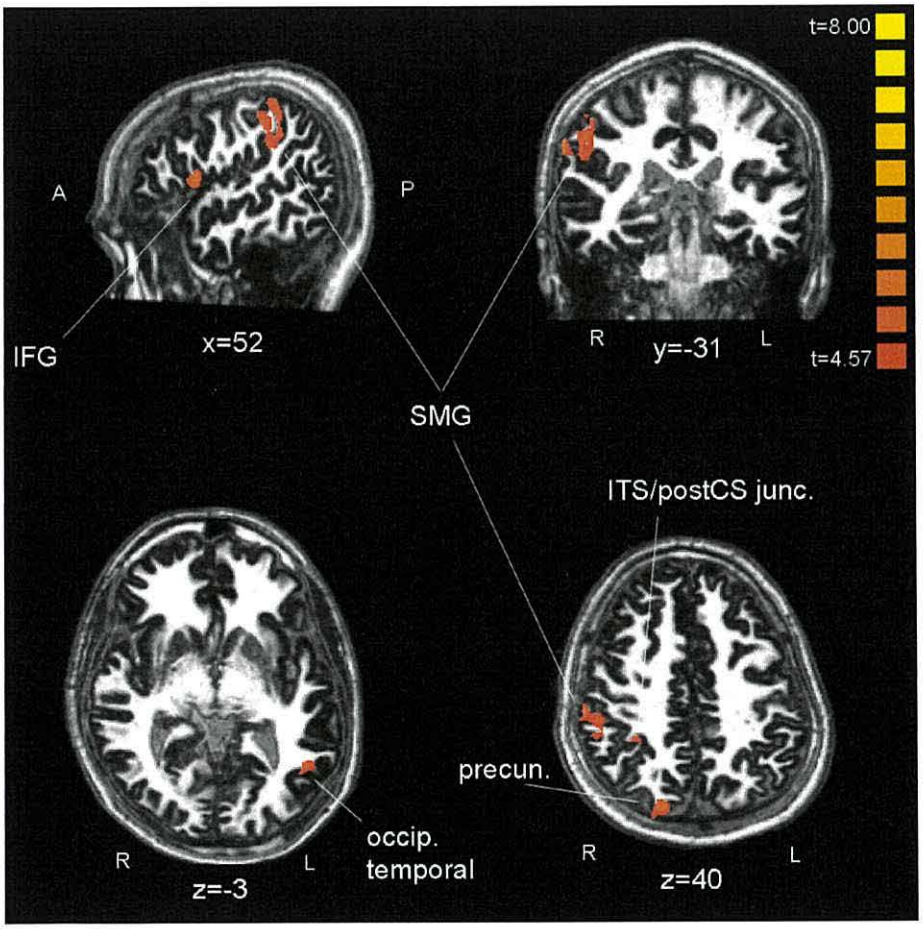


Figure 3. Activations from the whole-brain analysis. Coloured regions are those which responded more to coherent than to incoherent sequences, in a whole-brain, fixed-effects group average analysis. For clarity these regions are overlaid on an anatomical scan from a single participant. Activations are thresholded at $p < 0.000005$, uncorrected. Only activations greater than 100 mm^3 are shown. See Table 1 for details.

We found no significant clusters greater than 100 mm^3 that responded more to incoherent than to coherent sequences. In an exploratory analysis, the contrast was repeated at a threshold of $p < 0.005$ uncorrected. Activations were observed in several regions, including right orbitofrontal and superior frontal areas, medial prefrontal cortex, right occipital and occipito-temporal cortex, and right inferior occipitotemporal cortex (Table 2). Given the absence of any coherence effect in the independently-localised MT (see above), we attribute the increased activity to the incoherent sequences seen in early visual cortex to a greater amount of high-contrast visual transients, relative to the coherent condition. The peak coordinates of the right occipitotemporal region (45, -73, -1) were similar to those found in the EBA localiser (48, -70, 0). To explore this similarity further, we extracted the response of this region, as defined in the group-average incoherent vs. coherent contrast, to each of the four stimulus categories tested in the EBA localiser experiment (bodies, faces, tools, and scenes). A one-way ANOVA on these values showed a significant effect of category, $F(3,45) = 9.0$, $p < 0.001$. The mean response to bodies ($M = 6.9$) was greater than to scenes ($M = 3.2$), faces ($M = 4.0$), and tools ($M = 4.4$), all $t(15) > 3.2$, all $p < 0.01$. Thus the preference for incoherent over coherent sequences in the right EBA (defined in individual ROIs) is sufficiently strong and consistent to be observed at the group-average level, albeit at a liberal statistical threshold. Additionally, this preference does not appear to spread extensively within extrastriate cortex.

Table 1. All activations for which coherent > incoherent, from a fixed-effects multiple regression analysis, thresholded at T > 4.57, p < 0.000005 uncorrected. Only activations above a minimum spatial extent of 100 mm^3 are shown. Each row gives the location of the cluster, the volume of the activation, the location of the peak voxel of that activation in Talairach coordinates, the maximum and mean T value for the region, and the associated *p* value for the region as a whole.

Coherent > Incoherent

Region	extent (mm^3)	Mean Peak X	Y	Z	Max (T)	Mean (T)	Mean (p)
R supramarginal gyrus	2187	56	-31	35	6.1	4.92	0.000002
R dorsal precuneus	1061	19	-75	41	6.27	4.96	0.000002
R IPS/postcentral junction	833	37	-44	51	5.71	4.93	0.000002
R ventral inferior frontal gyrus	427	52	8	7	6.22	5.08	0.000001
L ventral occipitotemporal	119	-50	-58	-4	5.34	4.79	0.000002

Table 2. All activations for which incoherent > coherent, from a fixed-effects multiple regression analysis, thresholded at T > 2.81, p < 0.005 uncorrected. (There were no significant activations for this contrast at the threshold of T > 4.57 reported in Table 1). Only activations above a minimum spatial extent of 100 mm^3 are shown.

Incoherent > Coherent

Region	extent (mm3)	Mean Peak X	Y	Z	Max (T)	Mean (T)	Mean (p)
L posterior occipital lobe	1324	-12	-96	-1	5.01	3.40	0.0014
R angular gyrus	631	47	-66	32	3.64	3.08	0.0024
R inferior occipitotemporal	390	45	-73	-1	3.67	3.06	0.0026
R posterior occipital lobe	380	14	-94	7	5.18	3.60	0.0010
R orbitofrontal cortex	341	17	42	-1	4.69	3.36	0.0020
R medial prefrontal	242	1	38	23	3.49	3.00	0.0030
R superior frontal gyrus	107	19	18	54	3.74	3.19	0.0020

DISCUSSION

Our aim for this study was to distinguish two hypotheses about the functional role of the EBA within the network of cortical areas that analyze the human form and its characteristic motions. We found that, in contrast to both pSTS and MT+ (localised in individual subjects), and frontal and parietal regions (identified in a whole-brain analysis), the right EBA responds significantly more to incoherent action sequences than to coherent sequences. This preference is particularly striking given previous evidence for an enhanced EBA response to point-light biological actions relative to scrambled controls (Downing et al., 2001b; Grossman et al., 2002). Our findings are inconsistent with the "dynamic representation" hypothesis. They indicate that the EBA's functional contribution to the perception of human action is distinct from that of the other relevant areas involved in action perception.

The EBA did not respond more to coherent than to incoherent sequences; in fact it showed a significant bias in the opposite direction. One possibility is that this is due to the presence of more motion in the transition between frames in that condition, compared to the coherent condition. This seems unlikely given that motion-selective area MT+ showed essentially identical responses to the coherent and incoherent sequences, a significantly different pattern from that seen in the EBA. We instead propose that the EBA effect represents a form of neural adaptation. Several studies have found reduced responses in visual cortex to repeated or "primed" stimuli, compared to previously unseen stimuli (e.g. Buckner et al., 1998). This effect has been used as a tool to reveal the

functional properties of cortical areas, by assessing the kinds of stimulus changes to which the adaptation effect (and presumably the neurons under investigation) are invariant (e.g. Grill-Spector and Malach, 2001; Kourtzi and Kanwisher, 2001; Vuilleumier, Henson, Driver, and Dolan, 2002). In the coherent condition of the present study, the configuration of the actor's limbs changed relatively little, on average, from frame to frame (see Fig. 1). A region that processes the static visual appearance[1] of the human figure would be expected to respond less in this situation, due to the similarity between succeeding stimuli, compared to a situation in which the configuration changed much more dramatically from frame to frame. This effect was not predicted for the EBA, so this account is necessarily post-hoc, and requires further tests. An open question is whether this region also processes the static structure of non-body object kinds. If this were the case, one would expect a similar effect of coherence for these other object kinds. However, this seems unlikely owing to a) the weak response to other categories in EBA (Downing et al. 2001; in press) and b) the finding that TMS over the EBA selectively impairs visual short term memory for bodies (Urgesi et al., 2004).

In common with previous studies of fMRI adaptation effects (see Henson 2003 for a review), it is difficult to distinguish between an account in which neural activity is suppressed when identical or similar stimuli are perceived, and one in which activity is enhanced due to the novelty of changing stimuli. Likewise the effects of repetition on attentional state and neuronal adaptation can be difficult to disentangle. We note however that in the present study the reduced EBA activity to coherent sequences is unlikely to be

[1] Here we emphasize the EBA's sensitivity to changes in *configuration*, i.e. gross changes in the arrangment of the limbs relative to each other. We do not test the additional possibility that the EBA also represents the *form* of the body, e.g. differences in the contour and texture of body parts (cf. Urgesi et al., 2004).

due to global changes in arousal, given that on the whole, across the whole brain far more neural activity was elicited by the coherent relative to the incoherent sequences.

A secondary finding from the ROI analyses was that while the FBA did not significantly distinguish coherent from incoherent sequences, its response was similar to that found in the EBA (and this pattern significantly interacted with pSTS). Tentatively, we group the FBA, in terms of its functional properties, with EBA rather than pSTS. This raises the possibility that the observed adaptation effect originates in one of the two body-selective areas and is transmitted to the other (cf. Tolias et al., 2005). A key question for further studies is how the EBA and FBA interact. High temporal-resolution sampling could be used to test whether the functional connectivity between these regions is directionally biased. Additionally, manipulations of tasks and or stimuli could be tested in order to determine whether the EBA and FBA are functionally dissociable.

Finally, the whole-brain analyses confirmed, at the group level, the preference for incoherent sequences in the EBA (at a liberal statistical threshold). In contrast, several regions were activated more by coherent than incoherent sequences (see Figure 3 and Table 1). Some of these regions have been identified in previous neuroimaging studies of action observation. A classic study measuring brain responses to finger movements and to observation of similar movements (Iacoboni et al., 1999) found activity common to both tasks in the left frontal operculum (peak Talairach coordinates: -50, 12, 12) and at the junction of the right hemisphere intraparietal sulcus with the postcentral sulcus (peak: 37, -40, 57). We find similar activations, except that in the present study the frontal operculum activation is in the right hemisphere (operculum: 52, 8, 7; IPS/PostCS: 37, -

44, 51). Activations in these areas were also found in a recent fMRI study in which expert ballet and capoeira dancers observed movies of both ballet and capoeira dance moves (Calvo-Merino, Glaser, Grezes, Passingham, and Haggard, 2005). Effects of expertise (i.e. differential effects of observation for familiar vs unfamiliar actions) were found in the two regions mentioned above, as well as in the right supramarginal gyrus (peak MNI coordinates: 57, -30, 48), in line with the present findings (Talairach peak: 56, -31, 35). Thus our whole brain results support two main conclusions: 1) the coherent action sequences used as stimuli here activate areas that are known to be implicated in action observation, even though the motion signals typically present when viewing actions were absent; and 2) the EBA, as indicated by its preference for incoherent sequences, plays a different functional role than these other areas.

To summarise, we speculate that the EBA computes a static representation of the human body, and is not involved in analysis of biological motion *per se*. This is a somewhat paradoxical finding, in view of the EBA's enhanced response to intact vs scrambled point-light animations (Downing et al., 2001b; Grossman and Blake, 2002). This is in contrast to pSTS, which is strongly engaged by point-light animations, and weakly selective for coherent actions when the low level biological motion signals are removed. These findings are consistent with a recent computational model of biological motion processing (Giese and Poggio, 2003). In this model, subregions within a form-selective pathway (including the EBA) combine with separate motion-selective analysers to learn to recognise biological actions. The model proposes a set of "snapshot" neurons in the ventral stream that represent the various static postures that comprise an action.

The present findings suggest that the EBA contributes to this function, thus performing a fundamental step in our visual representation of the behaviour of other individuals.

These results provoke a number of questions for further research. First, it would be valuable to compare a wider variety of types of actions, particularly object-directed actions and semantically meaningful actions. Second, manipulations of subjects' tasks (e.g. in which the action itself is either task relevant or irrelevant) have previously been found to alter the brain activity elicited by viewing actions (e.g. Decety et al., 1997). Whether this would also modulate EBA activity is an open question, although the framework proposed here would predict no such effect, provided other factors such as attentional engagement were controlled. Finally, disruptive techniques could test the proposal that the EBA provides a crucial initial step in biological motion perception. Specifically, we predict that disruption of EBA activity by transcranial magnetic stimulation (TMS) would result in subsequent abnormal neural responses to biological motion in posterior superior temporal sulcus and fronto-parietal action systems.

Chapter 6:

Is the extrastriate body area involved in motor actions?

Astafiev *et al* (2004). report that unseen, visually-guided motor acts activate the extrastriate body area (EBA; Downing et al., 2001b). This finding has potential implications for understanding the interactions between motor and perceptual systems, and suggests a mechanism by which the visual stimulation resulting from one's own motor acts is distinguished from that produced by others (Jeannerod, 2004). We replicated Astafiev *et al.*'s experiment and found, in line with their findings, action-related modulation in EBA. However, a closer look showed that the region involved in visually guided motor acts is distinct from EBA, and that action-related modulation and body-selectivity are unrelated.

We scanned 13 subjects with an fMRI localiser for EBA (contrasting headless bodies with faces, scenes and tools). In the same session, we compared unseen visually-guided finger movements with a perceptually-matched control condition in an event-related design (see Figure 1). Replicating Astafiev *et al.*, we found a significant effect of finger movements in left ($t_{12} = 4.5$, $p < 0.001$) and right ($t_{12} = 4.0$, $p < 0.005$) EBA.

For each subject, a whole-brain contrast of finger movements versus control significantly ($p < 0.00001$, uncorrected) activated a bilateral temporal-occipital region (mean peak Talairach coordinates (x, y, z) left: -46, -65, -1; right: 53, -56, 0) that was close to EBA (left: -45, -74, -3; right: 48, -68, 0). The peak of this action-related region (ARR), however, was significantly anterior to EBA (left: $t_{12} = 5.4$, $p < 0.001$; right: $t_{12} = 5.9$, $p < 0.001$). Moreover, the spatial overlap (Rombouts et al., 1998) of ARR with EBA (at $p < 0.0005$, uncorrected) was only 14% (see Figure 2).

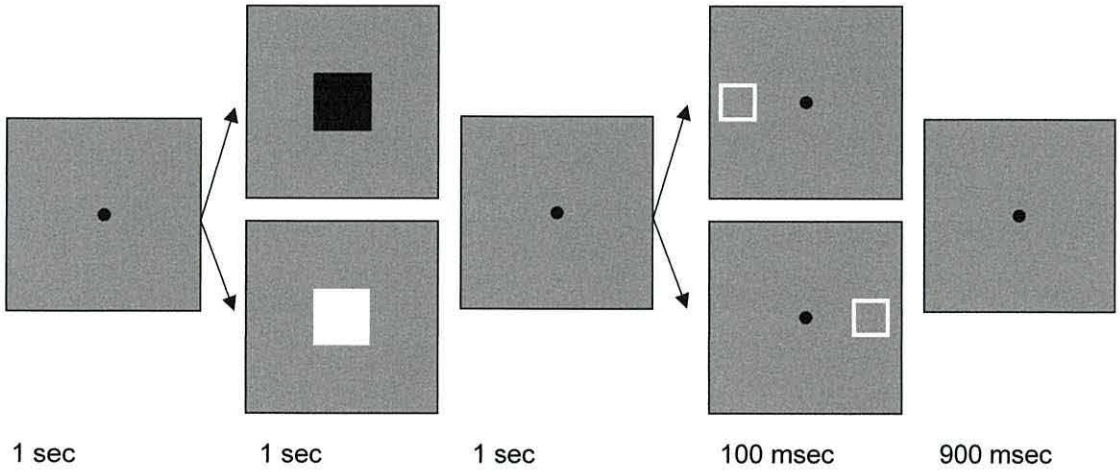


Figure 1. Schematic overview of the pointing task and the control condition. Subjects were presented with a black or white central square that indicated whether or not they had to move their finger in the direction of the subsequently presented target (consisting of a peripheral onset). For half of the subjects, the black square cued the movement condition and the white square the passive viewing condition; for the other half this assignment was reversed.

Note that the partial overlap of ARR and EBA does not necessarily mean that the same neurons are involved in both motor actions and body-perception. If this were the case, we would expect a positive voxel-by-voxel correlation between selectivity for bodies and action-related modulation. To test this we defined, for each subject, the intersection of ARR and EBA and calculated the correlation between the strength (as expressed by T values) of action-related activity compared to control, and body selectivity. The average correlation between these two measures was not statistically different from zero ($r = 0.00$, $p = 0.96$). This suggests that the region shared by ARR and EBA contains interleaved but functionally independent neural populations.

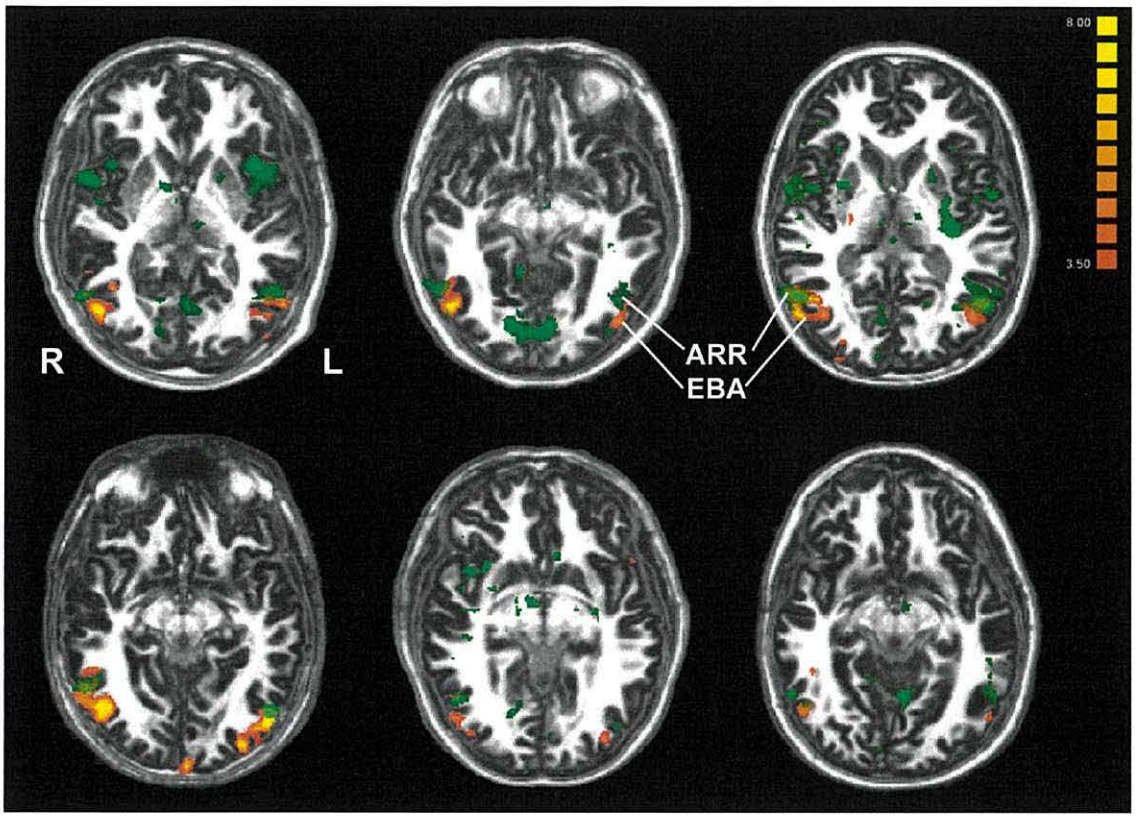


Figure 2. Axial slices showing EBA (orange) and ARR (green) in six subjects ($p < 0.0005$). In each subject, ARR is located anterior to EBA and overlaps only partly with it

To verify these findings, we scanned 5 subjects with an additional EBA localiser, using a contrast of body parts versus object parts. We replicated all of our key findings: a significant effect of pointing within the EBA ($p < 0.05$), a significantly anterior peak of ARR compared to EBA ($p < 0.05$), low spatial overlap of EBA and ARR (19%), and if anything a negative voxel-by-voxel correlation between action-related activity and body-selectivity ($r = -0.14$, $p = 0.08$). In contrast, the correlation between selectivity for whole bodies and for body parts was significantly positive ($r = 0.42$, $p < 0.05$), showing that the

absence of a correlation between action-related activity and body selectivity was not due to insufficient statistical power.

Thus the temporal-occipital area that is involved in executing motor actions is distinct from EBA. It may instead correspond to an area anterior to MT that is activated when subjects generate action-related words (Martin et al., 1995). It also falls near the putative human homologue of MST, which represents visual motion in the periphery (Dukelow et al., 2001; Huk et al., 2002). Further studies will be needed to determine the relationship between motor activity, action representation, and visual motion-selective regions in lateral temporal cortex.

Chapter 7:

General discussion

The studies presented in this thesis investigated the prevalence of body-selective areas in human visual cortex, the stability of these areas over time, and the involvement of these areas in biological motion perception, action perception, and the execution of motor actions.

In all experiments (and indeed in all 82 subjects tested) we replicated the finding of a bilateral body-selective area in posterior inferior temporal sulcus; the extrastriate body area (EBA; Downing et al., 2001a). The EBA was not only reliably activated across subjects, but also within subjects across time (Chapter 2). The location (in Talairach space) of the EBA changed very little across scanning sessions separated by 3 weeks. The small changes in location that were observed across scanning sessions were likely due to the separate coregistration of functional to structural data for the between-session comparison. One of the practical implications of this finding is that the EBA can be localized in a different session than the experimental session, for instance when time constraints do not allow localization in the same session. Other areas for which we assessed the within-subject reproducibility were the occipital face area (OFA), fusiform face area (FFA), and parahippocampal place area (PPA). Similar to the EBA, these areas showed little variability in their location over time. Functional reproducibility, as expressed by the stability of T-values across sessions, was also high for these areas.

Apart from the EBA, we consistently found another body-selective focus, located in the fusiform gyrus. In a group analysis, this region almost completely overlapped the face-selective FFA. An obvious explanation for this activation could thus be that it reflects an indirect activation of face-selective neurons, for instance through mental imagery or semantic association of (headless) bodies and faces. In this thesis I provide

evidence against this explanation and argue instead for the existence of a separate body-selective area, the fusiform body area (FBA). Several findings in this thesis support this claim: (1) The functional profile of the FFA and FBA are significantly different (Chapter 3); The FFA responded significantly more to faces than bodies, whereas no such difference was observed in the FBA. (2) The location of the activation peaks of the FFA and FBA were significantly different (Chapter 3). Interestingly, the direction (expressed by Talairach coordinates) of this location difference was not consistent across subjects, resulting in identical activation peaks when averaging the data from multiple subjects. (3) The FBA, but not the FFA, distinguished between schematic body representations (stick figures; Chapter 3). (4) The activation pattern across voxels in the fusiform gyrus was different for faces and bodies (Chapter 4). For example, body selectivity correlated with biological motion selectivity but face selectivity did not. In other words, voxels that were highly body selective were also highly biological-motion selective, whereas no such relation existed between face and biological-motion selectivity. In another analysis we showed that body selectivity, but not face selectivity, correlates significantly with selectivity for stick figures (unpublished data). In both these studies, the results were independent of the specific fusiform voxels that were selected for the analysis; the same results were found when selecting voxels in FFA, FBA, or both. This suggests that face- and body-selective neurons occupy the same cortical territory in the fusiform gyrus but are not identical.

The third main focus of this thesis is on the role of body-selective areas in biological motion perception, action perception, and the execution of motor actions. One important function of body-selective areas might be the coding and interpretation of other

people's actions. Given its close proximity to motion-selective area MT+, the EBA seems well-suited for encoding the dynamic aspects of bodies. Our results show that, even though EBA and FBA are activated by moving body stimuli and point-light displays of biological motion, the dynamic aspect of these displays is not crucial (Chapter 5). We attribute the response in the EBA to point-light displays to the form-from-motion information in these displays. In Chapter 6 we show that the EBA is not involved in the execution of motor actions, as suggested by Astafiev et al. (2004).

Below, I will discuss how these findings increase our understanding of the organization of the posterior fusiform gyrus and the role of the EBA in body perception. Finally, I will suggest possible directions for future research on these topics.

Organization of the posterior fusiform gyrus

Since the discovery of face selectivity in the right posterior/mid fusiform gyrus, and the subsequent proclamation of the "fusiform face area" (Kanwisher et al., 1997), the word fusiform gyrus has become almost synonymous with face processing and FFA. Indeed, many fMRI studies that report activity in the right fusiform gyrus label such activations "FFA", and forcedly try to link this activation to face processing. This includes studies that are not directly investigating face processing and often don't functionally localize FFA. One example is a study reporting fusiform activation to emotional (versus neutral) body postures (Hadjikhani and de Gelder, 2003). The authors of this study label their right fusiform activation "FFA", and conclude: "Our data suggest that the FFA and the amygdala, so far mostly associated with facial expression of fear in

the recent brain imaging literature, play a broader role than has been recognized so far." Based on the findings in Chapter 3, the increased fusiform response to fearful body postures in this study can be more parsimoniously explained by assuming that activity in body-selective visual regions can be "boosted" by emotional attention, as has been observed for faces (Vuilleumier et al., 2001; 2004). It is interesting to note that Hadjikhani and de Gelder did not report EBA activation for fearful body postures, suggesting that this region might be less susceptible to emotional influences than FBA.

Multiple distinct category-selective representations

Our finding of body selectivity in the fusiform gyrus (Chapter 3) led us to conclude: "These results challenge accounts of the mid-fusiform gyrus that focus solely on faces and suggest that this region contains multiple distinct category-selective neural representations." We proposed that future studies should scan at higher spatial resolution to expose these representations. This has now been done by Grill-Spector and colleagues (Grill-Spector et al., 2006), with striking results. They first defined the FFA in each subject using standard resolution (voxels of 3x3x3mm), and then examined the properties of this region using high spatial resolution (voxels of 1x1x1mm). They found localized clusters in the FFA that are highly selective to specific categories including non-face objects (animals, cars, and sculptures). More face- than object-selective voxels were found, explaining the overall face-selectivity of this region at standard resolution. Although bodies were not tested in this study, we might expect a relatively large number of body-selective voxels, given the strong overall response to bodies (e.g., Chapter 3).

What is the cause of this abundant category selectivity in this region, and which categories are likely to be represented? The evolutionary advantage for fast and accurate identification of the identity, actions, and emotional state of other people could explain the dedicated visual mechanisms for perceiving bodies and faces. Although this explanation might go some way in explaining the disproportionate body and face selectivity relative to other categories, it fails to explain the selectivity in the same region for modern categories such as cars. Thus, a complete explanation will have to include visual experience as another determinant of category selectivity. Evidence for selectivity as a function of experience comes from studies recording from neurons in monkey IT cortex (Logothetis et al., 1994; 1995). These studies show that after extensive training on recognizing novel shapes, some neurons start to respond selectively to these shapes. As mentioned in Chapter 1, fMRI studies on visual expertise have shown that expertise in a certain category can increase the response in the fusiform gyrus to this category (e.g., the response to cars in car experts). Although the authors of these studies argued for a domain-general, process-specific representation, it seems likely that scanning at higher resolution might reveal a fine-grained organization similar to that observed by Grill-Spector et al. (2006). For instance, the observation of increased activation to cars in car experts may reflect a larger number of, or an increased activation in, car-selective neurons. Interesting open questions concern the time scale at which these neurons tune to new categories, and whether there are critical periods in development for this tuning to occur. Apart from experience with a category, another potentially important factor in determining which categories are represented in the fusiform gyrus could be the similarity in configuration between members of a category, such that only categories

which members share a configuration definable by a fixed set of points are candidates for separate representation. Categories such as faces, bodies, and cars all have a highly constrained configuration. In contrast, categories such as vegetables and tools are more variable in their configuration, and are indeed less effective in activating the fusiform gyrus (Downing et al., in press). Future studies could test this notion in a more principled way, for instance by formally establishing the similarities between exemplars for different object categories while controlling for visual experience and expertise, or by devising new categories that vary systematically on these variables.

What is represented in the fusiform gyrus?

What is the nature of the representations in the fusiform gyrus? Selective activations are unlikely to reflect low level visual characteristics of the stimuli, as evidenced for instance by the selectivity in FBA for schematic body stimuli (stick figures; Chapter 3), and even for just a few moving dots that suggest a human figure (Chapter 4). Indeed, face-selective activation is also observed when subjects are merely imagining a face in the absence of any visual input (O'Craven and Kanwisher, 2000), or by faces that are occluded (Hulme and Zeki, in press). It is likely that other category-selective representations can similarly be activated without any visual input, and thus that the fine-grained organization of the fusiform gyrus is preserved for activations due to imagery. Furthermore, the same visual input can have different effects on activation in the fusiform gyrus depending on the subjective percept, for example in the face-vase illusion (Andrews et al., 2002). One possibility, then, is that the fusiform gyrus contains templates

for several categories that can be activated by bottom-up or top-down mechanisms, or by a combination of both, such that inputs that more closely resemble the template will be more likely to activate it. This template might contain a representation of the average of the experienced exemplars of a category. A recent monkey study, recording from anterior inferotemporal cortex, found that most of the face-selective neurons were tuned around an average face (Leopold et al., 2006). These findings support a norm-based recognition mechanism, in which faces are recognized by comparing them to an average face. Another study showed that faces may be represented by their direction (facial identity) and distance (distinctiveness) from a prototypical (mean) face (Loffler et al., 2005). This study made use of a multidimensional synthetic "face space" to systematically vary basic facial geometry (head shape, hair line, internal feature size and placement; Wilson et al., 2002). Different face identities were created by orthogonally varying the direction of the multidimensional vector in face space. Distinctiveness was varied by increasing the distance of this vector from the mean face. The results showed that the FFA represents faces by the direction and distance from a mean face, with increased activation to blocks of different faces with a large distance from the mean face. In contrast, FFA responses adapted when faces with the same identity but varying distinctiveness were presented, indicating an identity-related representation.

Finally, there is the possibility that neurons code for a particular exemplar of a category (so called "grand-mother cells"). Evidence for this type of sparse representation has been found for areas in the medial temporal lobe (Quiroga et al., 2005), but so far not for the fusiform gyrus.

The role of the EBA in body perception

The EBA responds generally and selectively to human bodies. Indeed, its response to bodies is high irrespective of image format, and no other object category tested so far activates this region to the same extent. What could be the function of this region in human cognition? A few hypotheses have been put forward in recent years, ranging from a social role involving the perception of other people's identity and actions, to more self-oriented functions such as the guidance of one's actions and the sense of agency. Below I will discuss these hypotheses in relation to the work presented in this thesis. I will argue that the current evidence suggests that the EBA is primarily involved in the simple detection, and perhaps identification, of the shape of human bodies and body parts. I will also put forward a new hypothesis concerning the type of objects that may be represented in EBA.

Identifying other people

One role of the EBA may be the identification of other people, perhaps under conditions in which face recognition is not possible (Downing et al., 2001a). We tested this hypothesis in a study where subjects were presented with images of bodies from different individuals holding varying poses. Subjects had to attend to either the identity or the posture of these images, and report whether two sequentially presented images where the same or different on the attended dimension. In individually localized EBA, we found strong activation for both tasks, and a small but significant preference for identity in right

(but not left) EBA (n=8; unpublished data). A whole brain analysis showed that the identity task (versus the posture task) activated regions that have previously been implicated in the perception of facial identity, including right posterior fusiform gyrus (Hoffman and Haxby, 2000) and medial prefrontal cortex (Mitchell et al., 2002). These results suggest that the right EBA, and possibly the right FBA, may be involved in perceiving the identity of other people. However, some caution in interpreting these results is warranted, as behavioral data suggested that subjects had more difficulty performing the identity task than the posture task. Also, it cannot be excluded that the left EBA is critically involved in both identity and posture discrimination. Future studies should use various difficulty-matched control tasks to test the involvement of right EBA and FBA in identity perception.

Tracking other people

Another function of the EBA could be the tracking of actions and intentions of other people, possibly in concert with mirror-neuron areas in frontal and parietal cortex. We tested this hypothesis in Chapter 5, where we manipulated the coherence of action sequences. We reasoned that if the EBA is involved in tracking actions of other people, it should be more activated by coherent (meaningful) compared to incoherent (scrambled) action sequences. As expected, the coherent sequences activated frontal (ventral premotor cortex) and parietal areas that have been implicated in action observation. In contrast to these areas, the EBA responded more to the incoherent sequences, suggesting that the EBA does not contain a dynamic, bound representation of actions. A recent TMS study

has confirmed this conclusion (Candidi et al., 2006). This study was aimed at testing the causal role of the EBA and ventral premotor cortex in processing body form and body action. Subjects had to discriminate either two different actions performed by the same body part, or the same action performed by two different models. TMS over EBA selectively impaired form discrimination, while TMS over ventral premotor cortex selectively impaired action discrimination (Candidi et al., 2006). Thus, these results provide more evidence for a functional dissociation between EBA and frontal mirror-neuron systems (ventral premotor cortex), with the former being involved in discriminating body form, and the latter being involved in discriminating body actions. Note that the images used in this study were static, and different actions were defined by different postures of the body part. As such, it is somewhat similar to the study described in the paragraph above, where we varied the task-relevance of posture and identity changes. Indeed, the results by Candidi et al. can be interpreted as reflecting a role for the EBA in detecting identity, given that differently formed body parts normally belong to different individuals. Future studies could try to distinguish body form and body identity by independently varying these factors. One way to do this is to create a synthetic "body space" similar to the "face space" created by Wilson et al. (2002) described above.

Guiding one's actions

Instead of being involved in the perception of other people, the EBA may be more concerned with the self. For instance, it could contain a visuo-spatial representation of one's own body parts, or it could be involved in guiding one's actions. To carry out these

functions, EBA would need to receive not only visual input but also motoric and proprioceptive signals. Some evidence for this notion has been provided by Astafiev et al. (2004). These authors claimed that the EBA is not only involved in the visual recognition of bodies but also in the execution of visually guided (unseen) movement of body parts (hands and feet). The authors suggest that such a movement may affect the actor's body representation through proprioceptive inputs that result from the movement. Alternatively, or additionally, it was suggested that the EBA could be activated through corollary discharge signals originating in motor areas. These signals may serve to dynamically update body representation in EBA, and adjust for sensory input resulting from the movement. Ultimately, these signals could function to distinguish between one's own and someone else's body (Jeannerod, 2004). Chapter 6 describes our attempt to replicate Astafiev et al.'s findings. We found, in line with their findings, action-related modulation in the EBA. However, a closer look showed that the region involved in visually guided motor acts only marginally overlapped with EBA. Furthermore, action-related modulation and body selectivity were unrelated on a voxel-by-voxel basis. That is, voxels that showed strong body selectivity did not necessarily show strong action-related modulation. We concluded that the temporo-occipital area that is involved in executing motor actions is distinct from the EBA.

The EBA codes for shapes that can change

Since its discovery in 2001, many studies have tried to uncover the precise role of the EBA. All of these studies have replicated the strong selectivity of the EBA for the

form of the human body and its parts. The EBA has also been shown to be necessary for accurate body recognition in a delayed match-to-sample task (Urgesi et al., 2004). Surprisingly, however, no study to date has provided strong evidence for a more specific or "high-level" role of the EBA in the processing of bodies, such as identity perception, action perception, or action guidance. A safe conclusion at this point is thus that the EBA is primarily driven by the shape of human bodies and body parts.

Here, I propose an account of the EBA that explains its selectivity profile by dynamic shape properties of objects. I start with assuming that EBA and the posterior/dorsal focus of LOC (LO) have the same function in the visual system, but differ in the objects for which they code. In this framework, LO is proposed to code for objects that change visual appearance mostly by rotational and translational movement, for instance through changes in the observer's viewpoint. EBA, in contrast, is proposed to code for objects that can actually change shape, e.g. through articulated movement of parts of the object. For example, the shape of a hand and the shape of a fist are very different, yet constitute the same object, and alternations between the two shapes can occur frequently.

The high degree of overlap between EBA and LO suggests that these areas are likely to be related, and are likely to share certain properties. If EBA is indeed analogous to area LO, many of the findings for LO could apply to EBA. For instance, LO shows sensitivity to location and size changes (Grill-Spector et al., 1999), but is cue invariant, that is, its response to a shape is independent of the features that define this shape (Kourtzi and Kanwisher, 2001). EBA's strong response to schematic body stimuli in Chapter 3 (stick figures) and 4 (moving dots) suggest that responses in EBA are similarly

cue invariant. Finally, LO has been shown to primarily code for the 2D shape of objects (Kourtzi et al., 2003), and responds preferentially to objects in the contralateral visual field (Niemeier et al., 2005). Future studies should test whether these results also apply to EBA, and could adopt adaptation paradigms similar to those used for LO (e.g., Grill-Spector, 1999; Kourtzi et al., 2003; Vuilleumier et al., 2002) to test for other similarities between these areas.

What computations are performed in EBA? One possibility is that it is involved in the transformation of shapes across configuration changes, in order to form a configuration-invariant representation that could underlie object identification. This configuration-invariant representation might reside either at the level of the EBA or in more anterior areas such as the FBA. A second possibility is that the representations in EBA (and LO) are used for the interaction with objects. Such representations are likely to be viewpoint and configuration dependent (James et al., 2002). These two possibilities correspond to two previously proposed types of representations in the visual system: viewer-centered representations (Bulthoff and Edelman, 1992) and object-centered representations (Biederman, 1987; Marr, 1982).

This framework can explain a few consistent findings. First, EBA does not respond strongly to a highly salient body part: faces (Downing et al., 2001a; 2006). Faces are an exceptional body part in that they are not able to move in an articulated fashion, and thus their 2D shape does not change with movement. As such, faces are more similar to other object types than to body parts. Second, EBA responds more to mammals, birds, and reptiles than to other object categories (Downing et al., 2006). Similar to human bodies, these categories change their shape frequently as a result of articulated

movements. Interestingly, the response in EBA to another animal category, fish, is low, consistent with the outlined framework. The finding by Downing et al. (2001a) of equal responses in EBA to unarticulated (e.g., hammer, screwdriver) and articulated objects (e.g., scissors, nutcracker) seems to contradict the hypothesis. Note, however, that these objects are very limited in the way they change shape through articulated movements compared to bodies and body parts. Thus, some degree of freedom in the movements may be required for the EBA to be activated.

How can this hypothesis about EBA be tested? One prediction is that the EBA responds significantly less to body parts that cannot move in an articulated manner, for example the chest or the back, compared to parts that can, such as hands and legs. A more controlled test would use newly devised object types and manipulate their variability in configuration across presentations. In a learning phase, exemplars of one of the types would be shown in different configurations and across different viewpoints, whereas the other type would be shown only across different viewpoints. Other (orthogonal) factors that can be included are the humanness of the shapes (e.g., containing legs and arms), the nature of the shape-changes (articulated versus non-articulated), and the degree of shape-changes. The prediction would be that after exposure to the two object types, subsequent exemplars of these types would activate EBA (and LO) to a different extent, with more activation in EBA to the type that was shown across different configurations.

In sum, the proposal views the EBA as an area that is analogous to LO. Similar characteristics are expected in both areas with respect to their sensitivity to changes in viewpoint, illumination, location, and size. The main difference between the two areas is the type of objects that they code for. Whereas LO codes for objects that change their

appearance mostly through rigid whole-body movements, EBA codes for objects that actually change shape, e.g., through articulated movements. Computations in EBA could facilitate the transformation of viewer-centered into object-centered representations, or could underlie our ability to interact with objects that change their shape.

Conclusions and future directions

Two areas in the human visual cortex are selectively activated by the perception of static and dynamic human bodies. The more posterior and dorsal EBA overlaps with object-selective area LO and motion selective area MT+, whereas the more anterior/ventral FBA overlaps with the ventral focus of object-selective LOC, as well as face-selective FFA. Despite the overlap between these areas, they are functionally separable (Chapter 3, 4, Downing et al., submitted).

What is the difference between EBA and FBA, and how do they interact? At present I can only speculate, as no study has shown differential responses in EBA and FBA or has directly investigated their interaction. A possibility is that the EBA is similar to LO and is involved in coding the 2D shape of bodies, and maybe other objects that share the dynamic shape properties of bodies, perhaps to allow successful interaction with these objects. The FBA might be similar to anterior/ventral LOC, and could be involved in the identification of bodies, for instance in combination with face-selective FFA. Future studies should systematically investigate these claims, and should test to what extent previous findings in posterior/dorsal and anterior/ventral parts of LOC apply to EBA and FBA, respectively. It is unknown, for example, to what extent the EBA and

FBA are sensitive to changes in viewpoint, illumination, location, and size of bodies. Future studies should also investigate whether EBA and FBA are differentially modulated by emotional and spatial attention (Hadjikhani and de Gelder, 2003; Vuilleumier et al., 2001; Wojciulik et al., 1998), and to what extent activity in EBA and FBA correlates with the successful identification and detection of bodies (e.g., Grill-Spector et al., 2004). Finally, functional connectivity measures such as Granger causality (Goebel et al., 2003) and dynamic causal modeling (Friston et al., 2003), as well as combined fMRI and EEG studies (e.g., Bledowski et al., 2006) could try to reveal the directionality of the interactions between EBA and FBA.

Chapter 4 and 6 demonstrated the power of multivariate pattern analysis in dissociating overlapping functional brain areas. This approach can be used to investigate many other questions where multiple areas overlap, not limited to visual cortex. For instance, future studies could investigate the similarities in activation patterns elicited by action perception and action observation in mirror-neuron areas.

To conclude, many areas in the brain are dedicated to perceiving and understanding other people's bodies. Posterior visual areas are likely involved in analyzing the shape of bodies, whereas more anterior visual areas might be involved in identifying these bodies. Regions in superior temporal sulcus have been shown to code for the actions and intentions of other people. Finally, parietal and frontal mirror-neuron areas may relate other people's bodily actions to our own actions, which may facilitate understanding and learning of these actions. Finding out how these areas interact to form the coherent and effortless percept of the identities, intentions and emotions of people around us is likely to remain a challenge for many years to come.

REFERENCES

Afraz SR, Kiani R, Esteky H. (2006) Microstimulation of inferotemporal cortex influences face categorization. Nature 442:692-695

Allison T, Ginter H, McCarthy G, Nobre AC, Puce A, Luby M, Spencer DD. (1994) Face recognition in human extrastriate cortex. Journal of Neurophysiology 71:821-825

Allison T, Puce A, Spencer DD, McCarthy G. (1999) Electrophysiological studies of human face perception. I: Potentials generated in occipitotemporal cortex by face and non-face stimuli. Cerebral Cortex 9:415-430

Allison T, Puce A, McCarthy G. (2000) Social perception from visual cues: role of the STS region. Trends in Cognitive Sciences 4:267-78

Andrews TJ, and Schluppeck D. (2004) Neural responses to Mooney images reveal a modular representation of faces in human visual cortex. Neuroimage 21:91-98

Andrews TJ, Schluppeck D, Homfray D, Matthews P, and Blakemore C. (2002) Activity in the fusiform gyrus predicts conscious perception of Rubin's vase-face illusion. Neuroimage 17:890-901

Astafiev SV, Stanley CM, Shulman GL, Corbetta M. (2004) Extrastriate body area in human occipital cortex responds to the performance of motor actions. Nature Neuroscience 7:542-8

Avidan G, Levy I, Hendler T, Zohary E, Malach R. (2003) Spatial vs. object specific attention in high-order visual areas. Neuroimage 19:308-18

Bartels A, and Zeki S. (2004) Functional brain mapping during free viewing of natural scenes. Human Brain Mapping 21:75-85

Beauchamp MS, Lee KE, Haxby JV, Martin A. (2002) Parallel visual motion processing streams for manipulable objects and human movements. Neuron 34:149-59

Behrmann M, Avidan G. (2005) Congenital prosopagnosia: face-blind from birth. Trends in Cognitive Sciences 9:180-7

Biederman I. (1987) Recognition-by-components: a theory of human image understanding. Psychol Rev 94:115–147

Bledowski C, Cohen Kadosh K, Wibral M, Rahm B, Bittner RA, Hoechstetter K, Scherg M, Maurer K, Goebel R, Linden DE. (2006) Mental chronometry of working memory retrieval: a combined functional magnetic resonance imaging and event-related potentials approach. Journal of Neuroscience 26:821-829

Bonda E, Petrides M, Ostry D, Evans A. (1996) Specific involvement of human parietal systems and the amygdala in the perception of biological motion. Journal of Neuroscience 16:3737-44

Boynton GM, Engel SA, Glover GH, Heeger DJ. (1996) Linear systems analysis of functional magnetic resonance imaging in human V1. Journal of Neuroscience 16:4207-21

Buccino G, Vogt S, Ritzl A, Fink GR, Zilles K, Freund HJ, Rizzolatti G. (2004) Neural circuits underlying imitation learning of hand actions: an event-related fMRI study. Neuron 42:323-34

Buckner RL, Goodman J, Burock M, Rotte M, Koutstaal W, Schacter D, Rosen B, Dale AM. (1998) Functional-anatomic correlates of object priming in humans revealed by rapid presentation event-related fMRI. Neuron 20:285-96

Bülthoff HH, Edelman S. (1992) Psychophysical support for a two-dimensional view

interpolation theory of object recognition. Proceedings of the National Acadamy of Science 89:60–64

Calvo-Merino B, Glaser DE, Grèzes J, Passingham RE, Haggard P. (2005) Action observation and acquired motor skills: an FMRI study with expert dancers. Cerebral Cortex 15:1243-9

Candidi M, Urgesi C, Ionta S, Aglioti SM. (2006) The causative role of EBA and vPMc in mapping body action and body form. An event-related repetitive Transcranial Magnetic Stimulation study. 12th annual meeting of the organization of Human Brain Mapping, Florence, Italy.

Carmel D, Bentin S. (2002) Domain specificity versus expertise: factors influencing distinct processing of faces. Cognition 83:1-29

Chan AW, Peelen MV, Downing PE. (2004) The effect of viewpoint on body representation in the extrastriate body area. Neuroreport 15:2407-10

Chao LL, Haxby JV, Martin A. (1999) Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. Nature Neuroscience 2:913-9

Chao LL, Martin A, and Haxby JV. (1999) Are face-responsive regions selective only for faces? Neuroreport 10:2945-2950

Cohen JD, and Tong F. Neuroscience. (2001) The face of controversy. Science 293:2405-2407

Cox D, Meyers E, and Sinha P. (2004) Contextually evoked object-specific responses in human visual cortex. Science 304:115-117

Cox DD, Savoy RL. (2003) Functional magnetic resonance imaging (fMRI) "brain

reading": detecting and classifying distributed patterns of fMRI activity in human visual cortex. Neuroimage 19:261-70

de Fockert JW, Rees G, Frith CD, Lavie N. (2001) The role of working memory in visual selective attention. Science 291:1803-6

Decety J, Grezes J, Costes N, Perani D, Jeannerod M, Procyk E, Grassi F, Fazio F. (1997). Brain activity during observation of actions. Influence of action content and subject's strategy. Brain 120:1763-77

Desimone R, Gross CG. (1979) Visual areas in the temporal cortex of the macaque. Brain Research 178:363-380

Desimone R, Albright TD, Gross CG, Bruce C. (1984) Stimulus-selective properties of inferior temporal neurons in the macaque. Journal of Neuroscience 4:2051-62

di Pellegrino G, Fadiga L, Fogassi L, Gallese V, Rizzolatti G. (1992) Understanding motor events: a neurophysiological study. Experimental Brain Research 91:176-80

Downing PE, Jiang Y, Shuman M, Kanwisher N. (2001a) A cortical area selective for visual processing of the human body. Science 293:2470-3

Downing P, Liu J, Kanwisher N. (2001b) Testing cognitive models of visual attention with fMRI and MEG. Neuropsychologia 39:1329-42

Downing PE, Bray D, Rogers J, and Childs C. (2004) Bodies capture attention when nothing is expected. Cognition 93:B27-B38

Downing PE, Chan A W-Y, Peelen MV, Dodds CM, Kanwisher N (in press). Domain specificity in visual cortex. Cerebral Cortex

Druzgal TJ, D'Esposito M. (2003) Dissecting contributions of prefrontal cortex and

fusiform face area to face working memory. Journal of Cognitive Neuroscience 15:771-84

Dukelow SP, DeSouza JF, Culham JC, van den Berg AV, Menon RS, Vilis T. (2001) Distinguishing subregions of the human MT+ complex using visual fields and pursuit eye movements. Journal of Neurophysiology 86, 1991-2000

Dumoulin SO, Bittar RG, Kabani NJ, Baker CL, Le Goualher G, Bruce Pike G, Evans AC. (2000) A new anatomical landmark for reliable identification of human area V5/MT: a quantitative analysis of sulcal patterning. Cerebral Cortex 10:454-63

Engel SA, Glover GH, Wandell BA. (1997) Retinotopic organization in human visual cortex and the spatial precision of functional MRI. Cerebral Cortex 7:181-192

Epstein R, DeYoe EA, Press DZ, Rosen AC, Kanwisher N. (2001) Neuropsychological evidence for a topographical learning mechanism in parahippocampal cortex. Cognitive Neuropsychology 18:481-508

Epstein R, Graham KS, Downing PE. (2003) Viewpoint-specific scene representations in human parahippocampal cortex. Neuron 37:865-76

Epstein R, Harris A, Stanley D, Kanwisher N. (1999) The parahippocampal place area: recognition, navigation, or encoding? Neuron 23:115-25

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

Fogassi L, Ferrari PF, Gesierich B, Rozzi S, Chersi F, Rizzolatti G. (2005) Parietal lobe: from action organization to intention understanding. Science 308:662-7

Friston KJ, Harrison L, Penny W. (2003) Dynamic causal modelling. NeuroImage 19:1273-12302

Gauthier I, Tarr MJ, Anderson AW, Skudlarski P, Gore JC. (1999) Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects. Nature Neuroscience 2:568-573

Gauthier I, and Logothetis N. (2000) Is face recognition not so unique after all? Cognitive Neuropsychology 17:125-142

Gauthier I, Skudlarski P, Gore JC, and Anderson AW. (2000) Expertise for cars and birds recruits brain areas involved in face recognition. Nature Neuroscience 3:191-197

Genovese CR, Lazar NA, Nichols T. (2002) Thresholding of statistical maps in functional neuroimaging using the false discovery rate. NeuroImage 15:870-878

Giese MA, Poggio T. (2003) Neural mechanisms for the recognition of biological movements. Nature Reviews Neuroscience 4:179-92

Goebel R, Roebroeck A, Kim DS, Formisano E. (2003) Investigating directed cortical interactions in time-resolved fMRI data using vector autoregressive modeling and Granger causality mapping. Magnetic Resonance Imaging 21:1251-1261

Grèzes J, Fonlupt P, Bertenthal B, Delon-Martin C, Segebarth C, Decety J. (2001) Does perception of biological motion rely on specific brain regions? Neuroimage 13:775-85

Grill-Spector K, Malach R. (2004) The human visual cortex. Annual Review of Neuroscience 27:649-677

Grill-Spector K, Kushnir T, Edelman S, Itzchak Y, Malach R. (1998) Cue-invariant activation in object-related areas of the human occipital lobe. Neuron 21:191–202

Grill-Spector K, Kushnir T, Edelman S, Avidan G, Itzchak Y, Malach R. (1999)

Differential processing of objects under various viewing conditions in the human lateral occipital complex Neuron 24:187–203

Grill-Spector K, Sayres R, Ress D. (in press) High-resolution imaging reveals highly selective nonface clusters in the fusiform face area. Nature Neuroscience

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

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

Gross CG, Rocha-Miranda CE, Bender DB. (1972) Visual properties of neurons in inferotemporal cortex of the Macaque. Journal of Neurophysiology 35:96-111

Grossman E, Donnelly M, Price R, Pickens D, Morgan V, Neighbor G, Blake R. (2000) Brain areas involved in perception of biological motion. Journal of Cognitive Neuroscience 12:711-20

Grossman ED, Blake R, Kim CY. (2004) Learning to see biological motion: brain activity parallels behavior. Journal of Cognitive Neuroscience 16:1669-79

Grossman ED, Blake R. (2001) Brain activity evoked by inverted and imagined biological motion. Vision Research 41:1475-82

Grossman ED, Blake R. (2002) Brain Areas Active during Visual Perception of Biological Motion. Neuron 35:1167-75

Hadjikhani N, de Gelder B. (2003) Seeing fearful body expressions activates the fusiform cortex and amygdala. Current Biology 13:2201-5

Halgren E, Dale AM, Sereno MI, Tootell RB, Marinkovic K, Rosen BR. (1999) Location

of human face-selective cortex with respect to retinotopic areas. Human Brain Mapping 7:29-37

Hasson U, Levy I, Behrmann M, Hendler T, Malach R. (2002) Eccentricity bias as an organizing principle for human high-order object areas. Neuron 34:479-490

Hasson U, Hendler T, Ben Bashat D, and Malach R. (2001) Vase or face? A neural correlate of shape-selective grouping processes in the human brain. Journal of Cognitive Neuroscience 13:744-753

Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P. (2001) Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science 293:2425-30

Haxby JV, Hoffman EA, Gobbini MI. (2002) Human neural systems for face recognition and social communication. Biological Psychiatry 51:59-67

Haynes JD, Rees G. (2005) Predicting the stream of consciousness from activity in human visual cortex. Current Biology 15:1301-7

Henson RN. (2003) Neuroimaging studies of priming. Progress in Neurobiology 70:53-81

Hoffman EA, Haxby JV. (2000) Distinct representations of eye gaze and identity in the distributed human neural system for face perception. Nature Neuroscience 3:80-84

Huk AC, Dougherty RF, Heeger DJ. (2002) Retinotopy and functional subdivision of human areas MT and MST. Journal of Neuroscience 22:7195-7205

Hulme OJ, Zeki S. (in press) The sightless view: neural correlates of occluded objects. Cerebral Cortex

Iacoboni M, Woods RP, Brass M, Bekkering H, Mazziotta JC, Rizzolatti G. (1999) Cortical mechanisms of human imitation. Science 286:2526-8

Ishai A, Haxby JV, Ungerleider LG. (2002) Visual imagery of famous faces: effects of memory and attention revealed by fMRI. Neuroimage 17:1729-41

Jackson PL, Meltzoff AN, Decety J. (2006) Neural circuits involved in imitation and perspective taking. Neuroimage 31:429-439

James TW, Humphrey K, Gati JS, Menon RS, Goodale MA. (2002) Differential effects of viewpoint on object-driven activation in dorsal and ventral streams. Neuron 35:793-801

Jeannerod M. (2004) Visual and action cues contribute to the self-other distinction. Nature Neuroscience 7:422-3

Jellema T, and Perrett DI. (2003) Cells in monkey STS responsive to articulated body motions and consequent static posture: a case of implied motion? Neuropsychologia 41:1728-1737

Jellema T, Perrett DI. (2003) Perceptual history influences neural responses to face and body postures. Journal of Cognitive Neuroscience 15:961-71

Johansson G. (1973) Visual perception of biological motion and a model for its analysis. Perception and Psychophysics 14:201-11

Kamitani Y, Tong F. (2005) Decoding the visual and subjective contents of the human brain. Nature Neuroscience 8:679-85

Kanwisher N, McDermott J, Chun MM. (1997) The fusiform face area: a module in human extrastriate cortex specialized for face perception. Journal of Neuroscience 17:4302-11

Kanwisher N, Stanley D, Harris A. (1999) The fusiform face area is selective for faces not animals. Neuroreport 10:183-7

Kanwisher N. (2000) Domain specificity in face perception. Nature Neuroscience 3:759-63

Kourtzi Z, Kanwisher N. (2000) Cortical regions involved in perceiving object shape. Journal of Neuroscience 20:3310-3318

Kourtzi Z, Bulthoff HH, Erb M, Grodd W. (2002) Object-selective responses in the human motion area MT/MST. Nature Neuroscience 5:17-18

Kourtzi Z, Erb M, Grodd W, Bulthoff HH (2003) Representation of the perceived 3-D object shape in the human lateral occipital complex. Cerebral Cortex 13:911-920

Kourtzi Z, Kanwisher N. (2001) Representation of perceived object shape by the human lateral occipital complex. Science 293:1506-9

Leopold DA, Bondar IV, Giese MA. (2006) Norm-based face encoding by single neurons in the monkey inferotemporal cortex. Nature 442:572-575

Levy I, Hasson U, Avidan G, Hendler T, and Malach R. (2001) Center-periphery organization of human object areas. Nature Neuroscience 4:533-539

Loffler G, Yourganov G, Wilkinson F, Wilson HR. (2005) fMRI evidence for the neural representation of faces. Nature Neuroscience 8:1386-1390

Loftus R, Masson EJ. (1994) Using confidence Intervals in within-subject designs. Psychonomic Bulletin and Review 1:476-90

Logothetis NK, Pauls J, Bülthoff HH, Poggio T (1994) View-dependent object recognition by monkeys. Current Biology 4:401–414.

Logothetis NK, Pauls J, Poggio T (1995) Shape representation in the inferior temporal

cortex of monkeys. Current Biology 5:552–563.

Logothetis NK, Pauls J, Augath M, Trinath T, Oeltermann A. (2001) Neurophysiological investigation of the basis of the fMRI signal. Nature 412:150-157

Machielsen WC, Rombouts SA, Barkhof F, Scheltens P, Witter MP. (2000) FMRI of visual encoding: reproducibility of activation. Human Brain Mapping 9:156-64

Malach R, Reppas JB, Benson RR, Kwong KK, Jiang H, Kennedy WA, Ledden PJ, Brady TJ, Rosen BR, Tootell RB. (1995) Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. Proceedings of the National Acadamy of Science 92:8135-9

Marr D. (1982) Vision. New York, Freeman Publishers

Martin, A., Haxby, J.V., Lalonde, F.M., Wiggs, C.L., Ungerleider, L.G. (1995) Science 270, 102-105

Michels L, Lappe M, Vaina LM. (2005) Visual areas involved in the perception of human movement from dynamic form analysis. Neuroreport 16:1037-41

Miki A, Liu GT, Englander SA, Raz J, van Erp TG, Modestino EJ, Liu CJ, Haselgrove JC. (2001a) Reproducibility of visual activation during checkerboard stimulation in functional magnetic resonance imaging at 4 Tesla. Japanese Journal of Ophthalmology 45:151-5

Miki A, Raz J, Englander SA, Butler NS, van Erp TG, Haselgrove JC, Liu GT. (2001b) Reproducibility of visual activation in functional magnetic resonance imaging at very high field strength (4 Tesla). Japanese Journal of Ophthalmology 45:1-4

Miki A, Raz J, van Erp TG, Liu CS, Haselgrove JC, Liu GT. (2000) Reproducibility of

visual activation in functional MR imaging and effects of postprocessing. American Journal of Neuroradiology 21:910-5

Miller MB, Van Horn JD, Wolford GL, Handy TC, Valsangkar-Smyth M, Inati S, Grafton S, Gazzaniga MS. (2002) Extensive individual differences in brain activations associated with episodic retrieval are reliable over time. Journal of Cognitive Neuroscience 14:1200-14

Mitchell JP, Heatherton TF, Macrae CN. (2002) Distinct neural systems subserve person and object knowledge. Proceedings of the National Acadamy of Science 99:15238-43

Mukamel R, Gelbard H, Arieli A, Hasson U, Fried I, Malach R. (2005) Coupling between neuronal firing, field potentials, and FMRI in human auditory cortex. Science 309:951-954

Niemeier M, Goltz HC, Kuchinad A, Tweed DB, Vilis T. (2005) A contralateral preference in the lateral occipital area: sensory and attentional mechanisms. Cerebral Cortex 15:325-331

Niessing J, Ebisch B, Schmidt KE, Niessing M, Singer W, Galuske RA. (2005) Hemodynamic signals correlate tightly with synchronized gamma oscillations. Science 309:948-951

Noll DC, Genovese CR, Nystrom LE, Vazquez AL, Forman SD, Eddy WF, Cohen JD. (1997) Estimating test-retest reliability in functional MR imaging. II: Application to motor and cognitive activation studies. Magnetic Resonance in Medicine 38:508-17

O'Craven KM, Downing PE, Kanwisher N. (1999) fMRI evidence for objects as the units

of attentional selection. Nature 401:584-7

O'Craven KM, Kanwisher N. (2000) Mental imagery of faces and places activates corresponding stimulus-specific brain regions. Journal of Cognitive Neuroscience 12:1013-23

Peelen MV, Downing PE. (2005a) Selectivity for the human body in the fusiform gyrus. Journal of Neurophysiology 93:603-8

Peelen MV, Downing PE. (2005b) Within-subject reproducibility of category-specific visual activation with functional MRI. Human Brain Mapping 25:402-8

Peelen MV, Downing PE. (2005c) Is the extrastriate body area involved in motor actions? Nature Neuroscience 8:125

Pelphrey KA, Mitchell TV, McKeown MJ, Goldstein J, Allison T, McCarthy G. (2003) Brain activity evoked by the perception of human walking: controlling for meaningful coherent motion. Journal of Neuroscience 23:6819-25

Pelphrey KA, Morris JP, Michelich CR, Allison T, McCarthy G. (2005) Functional Anatomy of Biological Motion Perception in Posterior Temporal Cortex: An fMRI Study of Eye, Mouth and Hand Movements. Cerebral Cortex 15:1866-1876

Perrett DI, Rolls ET, and Caan W. (1982) Visual neurones responsive to faces in the monkey temporal cortex. Experimental Brain Research 47:329-342

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. Behavioral Brain Research 16:153-70

Perrett DI, Hietanen JK, Oram MW, Benson PJ. (1992) Organization and functions of

cells responsive to faces in the temporal cortex. Philos Trans R Soc Lond B Biol Sci. 335:23-30

Peuskens H, Vanrie J, Verfaillie K, Orban GA. (2005) Specificity of regions processing biological motion. European Journal of Neuroscience 21:2864-75

Pinsk MA, DeSimone K, Moore T, Gross CG, Kastner S. (2005) Representations of faces and body parts in macaque temporal cortex: a functional MRI study. Proceedings of the National Acadamy of Science 102:6996-7001

Pruessmann KP, Weiger M, Scheidegger MB, Boesiger P. (1999) SENSE: sensitivity encoding for fast MRI. Magnetic Resonance in Medicine 42:952-62

Puce A, Allison T, Asgari M, Gore JC, McCarthy G. (1996) Differential sensitivity of human visual cortex to faces, letterstrings, and textures: a functional magnetic resonance imaging study. Journal of Neuroscience 16:5205-15

Puce A, Allison T, Gore JC, and McCarthy G. (1995) Face-sensitive regions in human extrastriate cortex studied by functional MRI. Journal of Neurophysiology 74:1192-9

Quiroga RQ, Reddy L, Kreiman G, Koch C, Fried I. (2005) Invariant visual representation by single neurons in the human brain. Nature 435:1102-1107

Ramsey NF, Kirkby BS, Van Gelderen P, Berman KF, Duyn JH, Frank JA, Mattay VS, Van Horn JD, Esposito G, Moonen CT and others. (1996) Functional mapping of human sensorimotor cortex with 3D BOLD fMRI correlates highly with H2(15)O PET rCBF. J Cerebral Blood Flow and Metabolism 16:755-64

Reed CL, Stone VE, Bozova S, Tanaka J. (2003) The body-inversion effect. Psychological Science 14:302-8

Rizzolatti G, Craighero L. (2004) The mirror-neuron system. Annual Reviews in Neuroscience 27:169-92

Rolls ET. (1984) Neurons in the cortex of the temporal lobe and in the amygdala of the monkey with responses selective for faces. Human Neurobiology 3:209-222

Rombouts SA, Barkhof F, Hoogenraad FG, Sprenger M, Scheltens P. (1998) Within-subject reproducibility of visual activation patterns with functional magnetic resonance imaging using multislice echo planar imaging. Magnetic Resonance Imaging 16:105-13

Rossion B, Schiltz C, Crommelinck M. (2003) The functionally defined right occipital and fusiform "face areas" discriminate novel from visually familiar faces. Neuroimage 19:877-83

Santi A, Servos P, Vatikiotis-Bateson E, Kuratate T, Munhall K. (2003) Perceiving biological motion: dissociating visible speech from walking. Journal of Cognitive Neuroscience 15:800-9

Saxe R, Jamal N, Powell L. (2006) My body or yours? The effect of visual perspective on cortical body representations. Cerebral Cortex 16:178-182

Saygin AP, Wilson SM, Hagler DJ, Bates E, Sereno MI. (2004) Point-light biological motion perception activates human premotor cortex. Journal of Neuroscience 24:6181-8

Schwarzlose R, Baker C, Kanwisher N. (2005) Separate face and body selectivity on the fusiform gyrus. Journal of Neuroscience 25:11055-9

Shiffrar M, Freyd J. (1990) Apparent motion of the human body. Psychological Science, 1:257-264

Spiridon M, Kanwisher N. (2002) How distributed is visual category information in human occipito-temporal cortex? An fMRI study. Neuron 35:1157-65

Stekelenburg J, and de Gelder B. (2004) The neural correlates of perceiving human bodies: an ERP study on the body-inversion effect. Neuroreport 15:777-780

Swallow KM, Braver TS, Snyder AZ, Speer NK, Zacks JM. (2003) Reliability of functional localization using fMRI. Neuroimage 20:1561-77

Talairach J, Tournoux P. (1988) Co-Planar Stereotaxic Atlas of the Human Brain. Thieme, Stuttgart

Tarr MJ, Gauthier I. (2000) FFA: a flexible fusiform area for subordinate-level visual processing automatized by expertise. Nature Neuroscience 3:764-9

Tegeler C, Strother SC, Anderson JR, Kim SG. (1999) Reproducibility of BOLD-based functional MRI obtained at 4 T. Human Brain Mapping 7:267-83

Tolias AS, Keliris GA, Smirnakis SM, Logothetis NK. (2005) Neurons in macaque area V4 acquire directional tuning after adaptation to motion stimuli. Nature Neuroscience 8:591-3

Tong F, Nakayama K, Moscovitch M, Weinrib O, Kanwisher N. (2000) Response properties of the human fusiform face area. Cognitive Neuropsychology 17:257-79

Tong F, Nakayama K, Vaughan JT, and Kanwisher N. (1998) Binocular rivalry and visual awareness in human extrastriate cortex. Neuron 21:753-759

Tootell RB, Reppas JB, Kwong KK, Malach R, Born RT, Brady TJ, Rosen BR, Belliveau JW. (1995) Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. Journal of Neuroscience 15:3215-30

Tsao DY, Freiwald WA, Knutsen TA, Mandeville JB, and Tootell RB. (2003) Faces and objects in macaque cerebral cortex. Nature Neuroscience 6:989-995

Tsao DY, Freiwald WA, Tootell RB, Livingstone MS. (2006) A cortical region consisting entirely of face-selective cells. Science 311:670-674

Urgesi C, Berlucchi G, Aglioti SM. (2004) Magnetic stimulation of extrastriate body area impairs visual processing of nonfacial body parts. Current Biology 14:2130-4

Vaina LM, Solomon J, Chowdhury S, Sinha P, Belliveau JW. (2001) Functional neuroanatomy of biological motion perception in humans. Proceedings of the National Acadamy of Science 98:11656-61

van Turennout M, Ellmore T, Martin A. (2000) Long-lasting cortical plasticity in the object naming system. Nature Neuroscience. 3:1329-34

Vuilleumier P, Armony JL, Driver J, Dolan RJ. (2001) Effects of attention and emotion on face processing in the human brain: an event-related fMRI study. Neuron 30:829-841

Vuilleumier P, Henson RN, Driver J, Dolan RJ. (2002) Multiple levels of visual object constancy revealed by event-related fMRI of repetition priming. Nature Neuroscience 5:491-9

Vuilleumier P, Richardson MP, Armony JL, Driver J, Dolan RJ. (2004) Distant influences of amygdala lesion on visual cortical activation during emotional face processing. Nature Neuroscience 7:1271-1278

Wachsmuth E, Oram MW, Perrett DI. (1994) Recognition of objects and their component parts: responses of single units in the temporal cortex of the macaque. Cerebral Cortex 4:509-22

Wilson HR, Loffler G, Wilkinson F. (2002) Synthetic faces, face cubes, and the geometry of face space. Vision Research 42:2909-2923

Wojciulik E, Kanwisher N, Driver J. (1998) Covert visual attention modulates face-specific activity in the human fusiform gyrus: fMRI study. Journal of Neurophysiology 79:1574-8

Yantis S, Serences JT. (2003) Cortical mechanisms of space-based and object-based attentional control. Current Opinion in Neurobiology 13:187-93

Yetkin FZ, McAuliffe TL, Cox R, Haughton VM. (1996) Test-retest precision of functional MR in sensory and motor task activation. American Journal of Neuroradiology 17:95-8

Yin RK. (1969) Looking at upside-down faces. Journal of Experimental Psychology 81:141-145