The Neural Representation of Facial Expression and Viewpoint in the Human Brain

Tessa Rosemary Flack

Doctor of Philosophy

University of York

Psychology

May 2016

Abstract

The human face conveys a wealth of information to the perceiver. Not only can we make judgments on the identity of the person, but we can also make finer interpretations about the emotional state of the individual, and what they are currently attending to. The ability to process and act upon this information effectively, facilitates successful social interactions. The key information that indicates to us how a person is feeling and what they are attending to, is their facial expression and facial viewpoint. Because of their dynamic properties, facial expression and viewpoint are described as changeable aspects of faces.

The human brain contains a core and extended network of face-responsive regions. One region in the posterior superior temporal sulcus (pSTS) is thought to have a central role in the processing of these changeable aspects of faces. An outstanding question about the neural representation of faces is whether changeable aspects such as expression and viewpoint, have distinct representations within these regions, or whether there are overlapping representations. This thesis aimed to further the understanding of the neural representation of facial expression and facial viewpoint, focussing on the neural representation in the pSTS.

First, this thesis investigated how expression is represented in the pSTS. A variety of behavioural evidence has shown that face images (in contrast to other objects) are processed holistically. In contrast to these findings, this chapter demonstrated that the pSTS represents facial expressions in a feature-based way, showing changes in response to any change in facial expression. However this chapter was also able to demonstrate that a region considered to be part of the extended face processing system, the inferior frontal gyrus (IFG), has a holistic representation of facial expression that reflects behavioural holistic processing.

The second experimental chapter asked whether there are distinct neural representations for processing changes in expression and changes in viewpoint, across the core and extended face-responsive regions. Distinct patterns of response were found for

changes in expression and viewpoint in the core regions. These representations were largely invariant to changes in identity, supporting the idea of distinct processing pathways for invariant and changeable aspects of faces. In contrast to the core regions, regions of the extended system (IFG and amygdala) were predominantly selective to changes in facial expression rather than changes in viewpoint.

The third experimental chapter asked whether there was a more fine-grained representation of facial viewpoint. Distinct patterns of response to different viewpoints were found. Interestingly, there was also a strong similarity for symmetrical viewpoints in the fusiform face area (FFA) and pSTS. This similarity in the patterns of response to symmetric viewpoint directions suggests that these regions represent an intermediate step towards full viewpoint invariance.

The final experimental chapter aimed to determine the relative dominance of expression and viewpoint in the neural representation of the core regions. The patterns of response were largely dominated by viewpoint, perhaps reflecting a neural coding that is linked to the visual properties of the face. Overall, these findings support the existence of distinct patterns of response to expression and viewpoint, whereby these changeable aspects of faces are represented by an overall pattern across the core face-responsive regions, rather than as discreet modules.

List of Contents

List of Figures

List of Tables

Acknowledgments

I would firstly like to thank my supervisors, Professor Andy Young and Professor Tim Andrews. It has been an absolute privilege working with you. I can't quite believe that this has almost come to an end. It has been both the best and hardest 3+ years of my life. I can't thank you enough for the amazing opportunities you have given me. I have come a long way from working on the checkouts at Tesco, and I have you both to thank for that. You have helped me change my life.

I would like to thank the staff of the York Neuroimaging Centre, without whom I could not have done this. I know how privileged I have been, to have access to such amazing facilities on the door step. Thank you also to Rob Jenkins for being a supportive member of my TAP panel. I really value the time you gave me, when I've just needed to talk.

I would also like to thank the members of Face Lab for being a fantastic and supportive lab group, and for all the fun we had together. I'd like to give a special thanks to David Watson who, despite his protests, has been a great friend. I am so grateful for all the support you have given me, and the patience you have had when I felt like I had no idea what I was doing. I miss our long talks about everything from science, politics and religion, to all those things we looked up, where we really should have had safe search on.

Finally, I would like to make a very special acknowledgement to my wonderful partner Paul Elliott. Not only did York give me the most amazing MSc and PhD experience, it also brought me my soon-to-be husband. You have been the most amazing support. Thank you in particular, for all the times where I said, "I've just got to finish something, I'll be five minutes" and an hour and a half later, you were still waiting. Thank you for consoling me when I cried over my data and thank you for being so calm and collected when I stressed out about every presentation I've ever given. Thank you, you are my hero.

Author's Declaration

This thesis contains work completed by the author under the supervision of Professor Tim Andrews and Professor Andy Young.

The empirical work presented in Chapter 5 has been published as:

Flack, T.R., Andrews, T.J., Hymers, M., Al-Mosaiwi, M., Marsden, S.P., Strachan, J.W.A., Trakulpipat, C., Wang, L., Wu, T. and Young, A.W. (2015). Responses in the Right Posterior Superior Temporal Sulcus Show a Feature-Based Response to Facial Expression. Cortex, 69, 14-23.

The author contributed equally to the design, stimulus generation, methods, analysis and write-up of the experiments reported in Chapter 5.

The results from multiple empirical chapters have been presented at the following conferences:

Flack, T.R., Young, A.W. & Andrews, T.A. (2016). Mirror-symmetric responses to viewpoint do not lead to an advantage in the recognition of unfamiliar facial identities. *Paper presented at the Annual Meeting of the Experimental Psychology Society, London, UK.*

Flack, T.R., Young, A.W. & Andrews, T.A. (2015). Neural Representations of Expression and Viewpoint Information in the Temporal Cortex. *Journal of Vision, 15*(12), 433-433. *Poster presented at the Annual Meeting of the Vision Sciences Society, Florida, USA.*

Flack, T.R., Andrews, T.J., Hymers, M., Strachan, J.W. & Young, A.W. (2015). Responses in the Right Posterior Superior Temporal Sulcus Reveal a Feature-Based Rather than Holistic Representation of Facial Expression. *Poster presented at the Applied Face Meeting, York, UK.*

Flack, T.R., Watson, D.W., Harris, R.J., Hymers, M., Gouws, A., Young, A.W., & Andrews, T.A. (2014). Distinct Representations for Rigid and Non-Rigid Facial Movements in Face-Selective Regions of the Human Brain. *Paper presented at the Annual Scientific Meeting of the British Association for Cognitive Neuroscience, York, UK.*

Flack, T.R., Watson, D.W., Harris, R.J., Hymers, M., Gouws, A., Young, A.W., & Andrews, T.A. (2014). Distinct Representations for Rigid and Non-Rigid Facial Movements in Face-Selective Regions of the Human Brain. *Journal of Vision, 14*(10), 1383-1383. *Poster presented at the Annual Meeting of the Vision Sciences Society, Florida, USA.*

Flack, T.R., Young, A.W. & Andrews, T.J. (2013). Mapping the Functional Organization of the Face-Responsive Regions in the Human Brain using fMRI. *Poster presented at the Centre for Chronic Diseases and Disorders Conference, York, UK.*

All sources are acknowledged as References.

Chapter 1 – Literature Review

1.1 The human face

The human face plays a crucial role in social communication. The face conveys a huge variety of information to the observer, allowing you to determine key basic information such as an individual’s gender and age, but also allows you to make finer interpretations such as how the person is feeling and what they are attending to. Being able to accurately interpret this type of information enables us to form strong relationships with other humans and have successful social interactions.

Human faces (and those of most other animals) have evolved to have a very specific design. There are two eyes, horizontally aligned. These are positioned above a single centrally placed nose. Beneath the nose, is a centrally placed mouth. These features fulfil biological roles, enabling us to see, smell and eat food. The necessary design of these features has meant that faces, both human and otherwise, follow a similar overall template. This means that humans have had to develop a very fine-tuned system in order to detect small changes not only across faces, but within individual faces. In order to detect whether we know an individual or not, we must be able to distinguish their features from other faces. Humans are extremely adept at familiar face perception, and in addition, the majority of humans are very good at interpreting even small changes in an individual’s face, and this is likely to have contributed to the success of our species.

Detecting whether a face belongs to someone we already know or is that of a stranger, is a fundamental skill. It is something that the majority of people find effortless, however when you consider the between-person similarity of faces, and the within-person variability of a particular face, it is no mean feat. Not only can a person’s face vary due to the viewing angle you have, and the lighting in the environment, it can also vary as a result of variability to do with the person, for example makeup, expression, weight change, hairstyle and accessories. Given the difficulty of this task, it is of no surprise that there is a huge literature devoted to the perception of identity. Identity is considered to be an invariant aspect of faces – a person’s identity doesn’t change. However, in addition to invariant aspects of faces, there are also many changeable aspects of the face, and these will be the focus of this thesis.

1.2 The changeable aspects of faces

As discussed above, the human face has both invariant and changeable aspects. For example a person's identity doesn't change, but their facial expression and viewpoint change regularly. Models of face processing propose that changeable aspects of faces such as expression are considered to be processed independently of invariant aspects such as identity (Bruce & Young, 2012; Haxby, Hoffman, & Gobbini, 2000). The following section will briefly outline the proposed separate pathways for processing invariant and changeable aspects of faces, before focusing on two types of facial change – expression and viewpoint.

1.2.1 Separate pathways for invariant and changeable aspects

The Bruce and Young (1986) model of face processing proposes that once the initial structural encoding and visual analysis of a face has occurred, there is a functional separation in the processing of facial identity and facial expression (Figure 1.1).

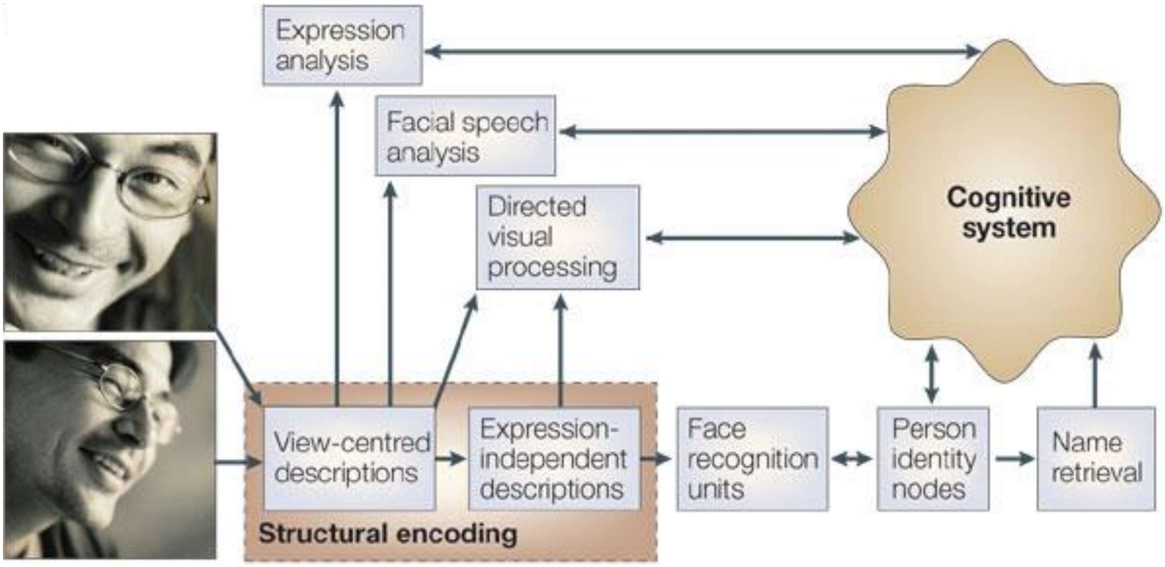


Figure 1.1. The Bruce and Young (1986) model of face processing, demonstrating separate routes for identity and expression recognition. Figure adapted from Calder & Young (2005), Box 1.

There is evidence to support the independent processing of identity and expression from neuropsychological cases studies that have shown that patients with prosopagnosia (impaired facial identity recognition) can have intact expression recognition (Baudouin & Humphreys, 2006; Tranel, Damasio & Damasio, 1988). In addition, a dissociation is seen where patients with impaired facial expression recognition have intact identity recognition (Humphreys, Donnelly, & Riddoch, 1993; Parry, Young, Saul & Moss, 1991; Young, Newcombe, de Haan, Small, & Hay, 1993). In addition, behavioural research has shown that processing of identity and expression can be performed relatively independently. For example Bruce (1986) and Young, McWeeny, Hay and Ellis (1986) both found that the familiarity of a face (famous or unfamiliar) did not have an effect on the judgement of facial expressions.

1.2.2 The role of changeable aspects of faces

The ways in which a human face changes are numerous and these changes convey different types of information to the observer. This information is particularly important for social communication. According to the Haxby et al. (2000) model, changeable aspects of faces include facial expressions, eye gaze and lip movements. However traditionally, movements of the face have been split into two categories; rigid and non-rigid movements (Bruce & Young, 2012). Rigid movements include movements of the head and eyes, and these types of movement do not change the shape of the facial features. Non-rigid movements however, involve movement of the facial muscles and therefore, movement of the facial features. Rigid movements of the head and eyes are often considered to convey submissiveness and dominance (Kleinke, 1986), but also signal the direction of attention. The interpretation of eye gaze in particular, is considered important for social interactions (Friesen & Kingstone, 1998). Non-rigid movements of the facial muscles give rise to facial expressions, which convey clear social signals and most expressions of emotion use a number of simultaneous muscle movements. These two types of movement (rigid and non-rigid) are considered to be changeable aspects of faces, however they are clearly different types of movements which convey different types of information to the observer. The remainder of this section will focus on the two changeable aspects of faces that will be the focus of this thesis; facial expression and viewpoint.

1.2.3 Facial expressions

Facial expressions are conveyed through a combination of changes in the facial muscles. Facial expression categories tend to be identified based on specific combinations of muscle changes. These changes give rise to a holistic perception that conveys a particular social signal (Calder, Young, Keane, & Dean, 2000). Darwin (1872) proposed that facial expressions evolved as signals which are based on a combination of muscular movements that have some kind of physiological benefit. For example, the expression of disgust involves scrunching up the nose which restricts the nasal passages, thus preventing you smelling an unpleasant odour. Darwin proposed that this association between emotional expressions and the underlying biologically adaptive function suggests that there are likely to be a number of universally recognisable expressions.

However, it is difficult to test whether facial expressions are universally recognised, as expression categories are so inherent in modern culture. Due to the wide reach of western culture through television and other media, it is hard to imagine being able to grow up without the standard cultural understanding of facial expression categories. At the time Darwin conducted his research though, this was much less of an issue and he collected a variety of evidence to support his theory. However, Darwin's theories on evolution of facial expressions were not well received at the time and his methods were criticised for being biased - his evidence tended to consist of observations made by himself and close colleagues (Darwin, 1872).

Ekman and his colleagues revisited this idea of universality many years later. This time, they addressed the issue of cultural influence by testing facial expression recognition across a number of cultures (Ekman, 1972; Ekman et al., 1987). These cultures ranged from all over the world, e.g. US, Germany, Scotland, Japan, Italy, Brazil, Argentina. Amongst these cultural groups, there was strong agreement on the interpretation and categorisation of facial expressions. If facial expressions are modified by cultural norms, then you might expect to find variation amongst cultures. However as mentioned earlier, western culture is wide-reaching and this could heavily influence the development in interpretation of facial expressions. All of the cultural groups mentioned above, could be considered to have a strong to moderate exposure to western culture through television and other media. To tackle this issue, Ekman and colleagues also tested a preliterate culture from New Guinea, with little if any contact with the outside world (Ekman, 1972). They presented participants with photographs of people posing facial expressions and a short story aimed to elicit a particular emotion or be associated with a particular emotion category e.g. ‘your child has died’. They were asked to pick the photographs that fitted the emotional category indicated by the short story. Accuracy was high, with anger, disgust, happiness and sadness all being distinguished from each other, and from fear and surprise. However fear and surprise were commonly confused. The New Guineans portrayals of these expressions were all accurately identified by American observers, although they again could not discriminate between their portrayals of fear and surprise (Ekman, 1972, 1980; Ekman et al., 1987).

Ekman (1992) proposed the term ‘basic emotions’ to refer to the emotions which are amongst other criteria, distinctive universal signals, have distinctive physiology and are present in other primates. These basic emotions are considered to be those that are universally recognised across the human species, including those in cultures that have had little interaction with the modern world, such as the New Guinea tribe described above. Although there are subtle variations across cultures in interpretations of some expressions, the following five (often termed the basic five) are generally agreed upon - fear, anger, sadness, happiness and disgust. There is some contention over whether surprise and contempt should be considered amongst this list. Surprise can be both negative and positive affect and is often confused with fear (Ekman, 1972) and there have been conflicting results regarding the recognition of contempt (Ekman & Friesen, 1986; Ekman & Heider, 1988; Elfenbein & Ambady, 2002; Izard & Haynes, 1988). As a result, the basic five are those most commonly used in studies of facial expression and will be those used for the experiments described in this thesis.

1.2.4 Facial viewpoint

As outlined previously, facial expressions are considered to be non-rigid movements of the face. Although study of facial expression has dominated the literature on changeable aspects of faces, rigid movements are equally interesting and have important implications both for survival and communication.

Different types of rigid movement convey crucial information about a person’s focus of attention - head direction (viewpoint) and eye gaze direction allow us to assess what a person is looking at or attending to, and their facial expression allows us to assess whether what they are looking at is of relevance to us. This information is crucial in situations where we may be at threat and unaware of the danger which another person perceives. Rigid movements of the head and eyes (viewpoint and eye gaze) are also considered to convey messages regarding personal traits, such as submissiveness and dominance and can also convey information regarding a person’s liking and attraction towards another person or object (Kleinke, 1986).

Viewpoint also plays a significant role in face recognition. When we know someone well, we are able to identify them from most viewpoints and often in poor lighting conditions. Burton, Jenkins and Schweinberger (2011) point out that the biggest variation in images of the same identity, tend to be variation that is not due to the face itself, but due to changes in the world and previous research suggests that once a face has been detected, amongst the earliest information to be extracted is facial viewpoint (Or & Wilson, 2010). Clearly viewpoint has an important role, however there are times when this within-person variability is not helpful to us. When we want to recognise someone, we need to be able to discount changes in viewpoint, ignoring this variability as it only adds noise to the goal of recognition (Booth & Rolls, 1998; Kourtzi, Erb, Grodd, & Bülthoff, 2003).

There is a large body of research looking at how rigid movements of the face (both in terms of viewpoint and eye gaze) modulate responses to facial expressions. This will be more widely discussed in section 1.3.5. regarding neuroimaging research into the amygdala, however this type of effect has also been demonstrated behaviourally. A recent study by Guo and Shaw (2015) demonstrated that perceived intensity of expressions is significantly affected by the viewpoint of the face. Profile views of the face significantly decreased the perceived intensity of the facial expression compared to frontal views of the face. This was the case for all facial expressions tested; happiness, sadness, anger, fear, disgust and surprise. Interestingly, viewpoint had little effect on the categorisation accuracy of the expressions, despite ¾ profile and full profile views decreasing the number of features visible to the observer.

To summarise, viewpoint plays an important role in face processing. It can convey important information about an individual, such as their focus of attention, and give us indications about social traits. However it also provides variation we need to discount when it comes to recognising familiar individuals and has important implications for how we interpret facial expressions.

1.3 Neural regions implicated in face processing

There are many regions of the human brain that are associated in some way with face perception. Various theoretical models have sought to explain the involvement of these many regions, and to try and pinpoint specific aspects of face perception to specific regions. The following section will describe a popular, dominant model of the neural representation of face perception which has influenced the regions selected in the following experimental chapters.

1.3.1 The Haxby Model

Haxby et al. (2000) proposed a distributed neural system for face perception. This includes a core system involving the inferior occipital gyri (the occipital face area (OFA)), the lateral fusiform gyrus (also known as the fusiform face area (FFA)) and the posterior superior temporal sulcus (pSTS). In addition, an extended system includes areas such as the intraparietal suclus (IPS), auditory cortex, amygdala, insula, limbic system and anterior temporal lobe (ATL), (Figure 1.2).

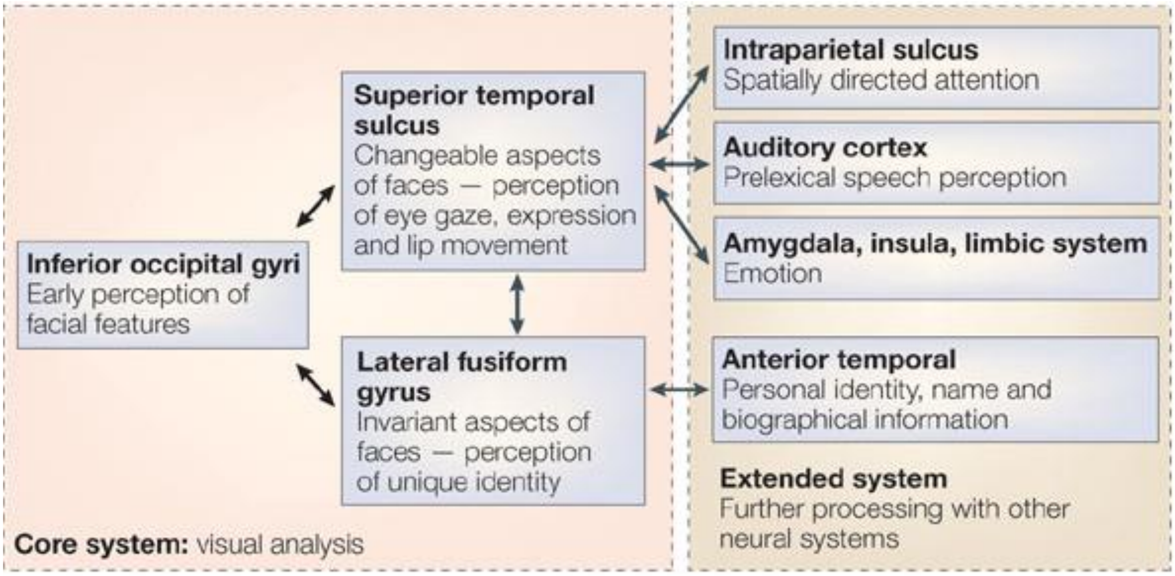


Figure 1.2. The distributed neural system for face perception as proposed by Haxby et al. (2000). Figure adapted from Calder & Young (2005), Box 1.

As discussed previously, the information available from a face can take the form of invariant information such as identity and gender, but also variant information such as viewpoint and facial expression. The Haxby et al. (2000) model identifies regions of the core system to be involved in the visual analysis of the face. The OFA is considered to be involved in the early perception of the facial features, such as the shape. The FFA is then thought to process more high level invariant aspects, for example unique identity. The posterior superior temporal sulcus (pSTS) is considered to be involved in the processing of the changeable aspects of faces. This information is then relayed to other regions in the extended face processing network, such as the amygdala and regions in the frontal and parietal lobes.

The Haxby et al. (2000) model shows a division into two processing pathways, emerging from the OFA. One pathway goes up to pSTS, and the other to the FFA. This separation in pathways, draws the line between the regions involved in the processing of the invariant information in faces, and the changeable aspects of faces. This separation in pathways is the neural representation of the division described above in section 1.2.1 (*Separate pathways for invariant and changeable aspects*).

The following section will give a brief overview of key regions identified by the Haxby model with a more detailed focus on those associated with the changeable aspects of faces.

1.3.2 The Occipital Face Area

The OFA is the first region implicated in neural models of face processing (Haxby et al., 2000; Ishai, 2008). Haxby and colleagues suggest the OFA is involved in the early perception of facial features and generates an initial representation of a face before further processing of aspects such as identity and expression occur. Rotshtein, Henson, Treves, Driver, and Dolan (2005) used a face morphing technique to demonstrate that the OFA releases from adaptation when changes to the physical appearance of a face are made, but crucially even when these changes do not result in the perception of a new identity.

Neurological literature has also implicated the OFA in deficits of face perception (Steeves et al., 2006). A meta-analysis looking at the associated lesions in 90 patients with prosopagnosia found the majority of prosopagnosia patients (and patients with achromatopsia with additional face-processing impairments) had lesions in the right OFA (Bouvier, 2005).

Research using transcranial magnetic stimulation (TMS) has demonstrated the OFA has a causal role in the early processing of faces (Pitcher, Walsh, Yovel, & Duchaine, 2007). Pitcher and colleagues found that TMS to the right OFA disrupted the discrimination of face parts. In a later study Pitcher, Charles, Devlin, Walsh and Duchaine (2009) were able to show that TMS to the right OFA significantly impaired participants in a face-matching task. In addition, a number of fMRI studies have implicated the OFA in the representation of parts of the face, including the eyes, nose and mouth (Liu, Harris, & Kanwisher, 2009; Nichols, Betts, & Wilson, 2010; Pitcher, Walsh, & Duchaine, 2011).

The evidence described above, fits with the hypothesis that the OFA is an initial component of the face perception network where it is involved in the early perception of facial features before subsequent processing occurs in further regions (Haxby et al., 2000).

1.3.3 The Fusiform Face Area

The FFA is a region thought to be selectively involved in the perception of faces (Kanwisher, McDermott, & Chun, 1997; Kanwisher & Yovel, 2006). The FFA is traditionally thought to be involved in the processing of invariant aspects of faces, such as facial identity (Grill-Spector, Knouf, & Kanwisher, 2004; Haxby et al., 2001; Hoffman & Haxby, 2000). Research by Andrews and Ewbank (2004) showed that responses to identity in the FFA reduced with repeated presentation of the same identity, but released from adaptation to images of different identities. Interestingly, adaptation to images of the same identity were invariant to changes in the size of the image. These results clearly implicate the FFA in the recognition of identity.

However, there have been a number of studies claiming to implicate the FFA in the processing of other aspects of faces as well as identity. Fox, Moon, Iaria, and Barton, (2009) found the FFA released from adaptation to changes in both identity and expression. This is supported by other studies showing sensitivity to facial expression in the FFA (Ganel, Valyear, Goshen-Gottstein, & Goodale, 2005; Narumoto, Okada, Sadato, Fukui, & Yonekura, 2001). There has also been some suggestion that responses in the FFA are more task-dependent rather than stimulus driven (Cohen Kadosh, Henson, Cohen Kadosh, Johnson, & Dick, 2010). Cohen Kadosh et al. found increased activity when expression changed during an identity task and increased activity when identity changed in an expression task. They also found increased activity in the anterior fusiform gyrus to changes in gaze in an expression task, supporting the role of the FFA in facial aspects other than identity. However they did not see increased activity to changes in gaze during an identity task.

There is also a contentious literature around the role of the FFA in expertise. Tarr and Gauthier (2000) argue the FFA is not specialised for faces, but for expertise. Faces being a category of objects where we have far greater expertise than others. Gauthier, Tarr, Anderson, Skudlarski and Gore (1999) demonstrated evidence of expertise recruiting the FFA with a set of novel objects (greebles) that participants were trained to become ‘experts’ with. The results have been fiercely debated not least because it is difficult to compare the results of expertise of greebles which required 7-10 hours of training (Gauthier & Tarr, 1997) with faces (a lifetime of experience), and that greebles themselves are inherently face-like (Figure 1.3), suggesting greeble expertise could reflect the recruitment of face-processing mechanisms (McKone & Kanwisher, 2005).

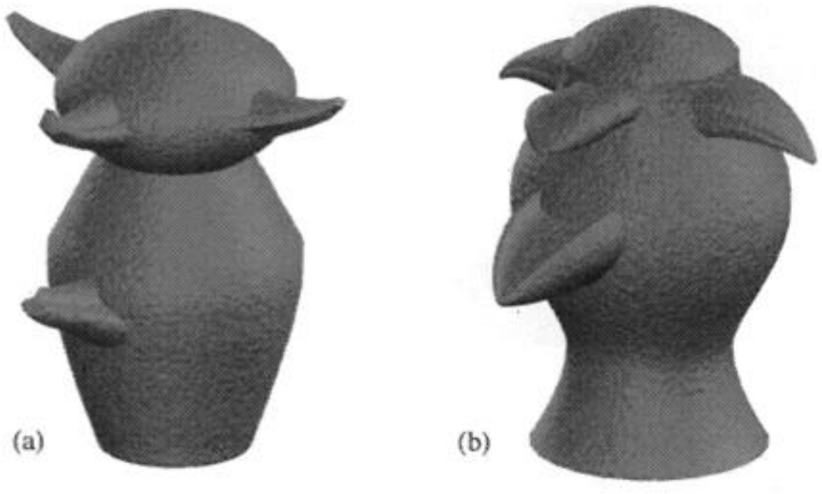


Figure 1.3. Examples of greebles used by Gauthier et al. in their expertise training. Greebles have two horizontally placed parts placed above two centrally placed parts – a kin to the positioning of the eyes, nose and mouth in faces. As pointed out by McKone and Kanwisher (2005), some look like they have heads (a) and some look like the entire greeble is a head (b). McKone and Kanwisher (2005), Figure 17.2.

In summary, the role of the FFA is highly debated. It is certainly a key area involved in face processing, however its role is thought by some to extend beyond only the invariant aspects of faces, to include changeable aspects such as facial expression.

1.3.4 The Superior Temporal Sulcus

The STS is widely considered to be responsive to facial expressions (Allison, Puce, & McCarthy, 2000; Calder & Young, 2005; Flack et al., 2015; Narumoto et al., 2001; Winston, Henson, Fine-Goulden, & Dolan, 2004) and neuroimaging studies have shown that the processing of viewpoint initially occurs in face-selective regions of the posterior superior temporal sulcus (pSTS) (Andrews & Ewbank, 2004). Work has largely focused on the right pSTS as significant activity is rarely seen in the left pSTS in standard face perception research using fMRI. The right pSTS can easily be found with face localiser scans when contrasting BOLD (blood oxygen level dependent) response to face stimuli with that of other objects. However, the left pSTS is localised much less frequently, suggesting the left pSTS may be performing a different role than that of the right pSTS.

It has been proposed that the pSTS is involved in processing socially relevant changes, and neuroimaging studies have supported this argument, demonstrating that the pSTS responds particularly strongly to changeable aspects of the same facial identity (Andrews & Ewbank, 2004, Harris et al., 2012) in comparison to different identities. This supports the idea that monitoring changes by a specific individual is often of particular social importance, for example, tracking the changes in a close friend or partner's emotional state. It has also been shown that the pSTS is not simply a face area, it is also responsive to movements of the eyes, mouth and hands (Pelphrey et al., 2005) and therefore appears involved in interpreting a range of biological movements.

Although predominantly considered to be involved in the processing of the non-rigid movements of facial expressions, the pSTS has also been shown to be involved in the processing of rigid movements of the head and eyes (Carlin & Calder, 2012; Fang, Murray, & He, 2007; Lee et al., 2010; Natu et al., 2010). A number of studies have also specifically implicated the posterior STS in the processing of gaze. Puce, Allison, Bentin, Gore and McCarthy (1998) used fMRI where participants viewed a face where the eyes were averted to the left or right. When the eyes altered between left and right aversion, and eyes directly looking at the observer, this activated the pSTS.

Some studies have found the pSTS is also more responsive to averted, rather direct gaze (Engell & Haxby, 2007; Hoffman & Haxby, 2000). However there is some inconsistency, with other studies finding stronger activation to direct gaze compared to averted gaze (Nummenmaa & Calder, 2009; Pelphrey, Viola, & McCarthy, 2004). Pelphrey et al. (2004) found stronger activation to mutual (direct) gaze, compared to averted gaze along the length of the STS. Pelphrey and colleagues used fMRI and virtual reality goggles to simulate a man walking towards the participant. The man either directed his gaze toward the participant, or away from them. This produced more activation for mutual gaze than averted gaze in the pSTS, however there was no such distinction seen in the fusiform gyrus, supporting the distinction between the roles of the pSTS and FFA in the Haxby et al., (2000) model.

It is possible however, that these results can be explained by the idea that the pSTS is affected by the context, rather than the direction of the gaze or head direction. Pelphrey et al. (2003) found that when an actor directs their gaze away from the target, the pSTS responds more strongly than when the actor directs their gaze towards the target. This suggests the STS may be involved in a higher level integration of the context and gaze.

As well as the visual responses described above, the pSTS can be seen as a multimodal region. It has been shown that the right pSTS is involved in multimodal integration of emotion (Hagan et al., 2009) where significant activity was found in the right pSTS to congruent audiovisual emotional stimuli. It has also been demonstrated that the left pSTS may be a multimodal region involved in the integration of vocal and facial speech signals (Calvert, 2001; Wright, Pelphrey, Allison, McKeown, & McCarthy, 2003).

As stated previously, much research has focused on the posterior region of the STS, however the location of activity in response to different types of biological movement (e.g. eyes, mouth, hands) seem to vary along the length of the STS (Figure 1.4), suggesting a relatively large area of cortex is involved in the processing of biological movements (Allison et al., 2000).

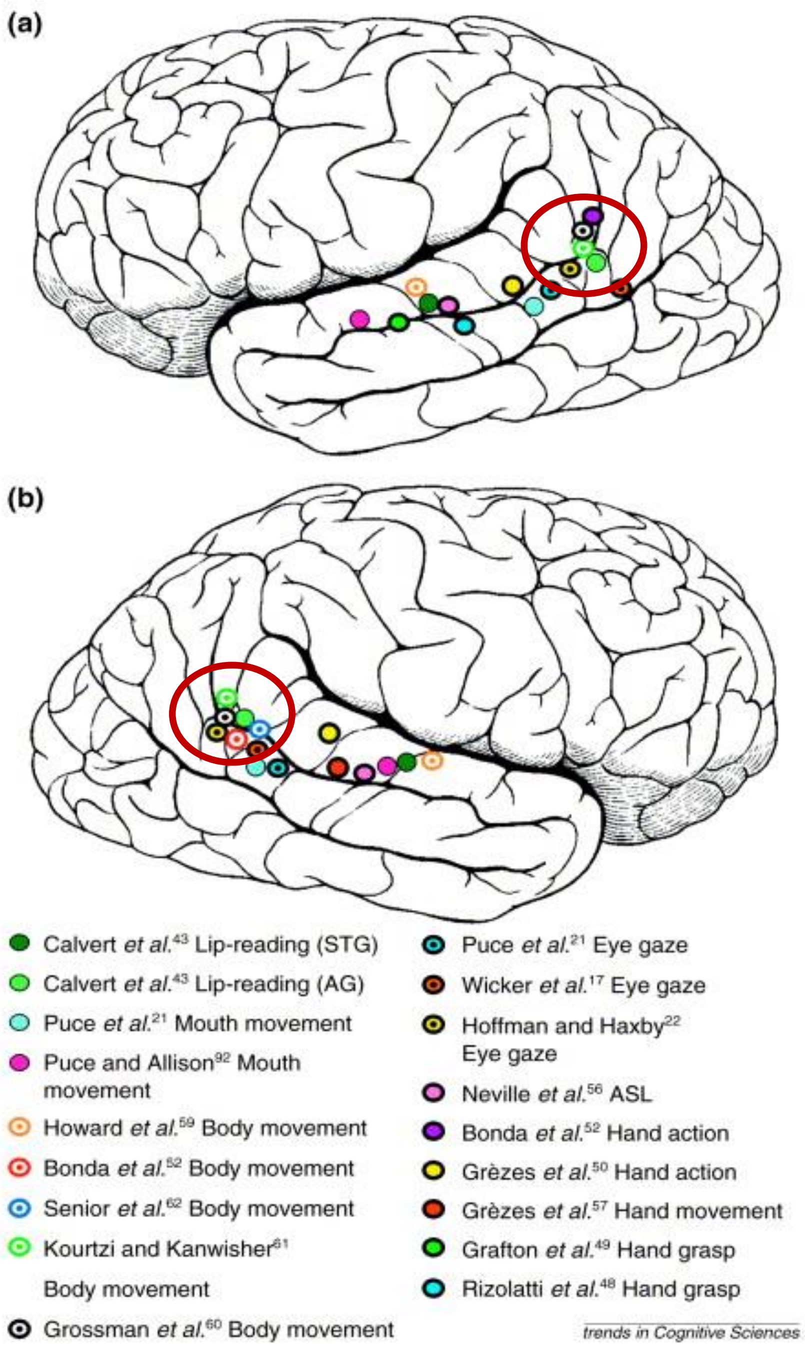


Figure 1.4. Variation along the length of the STS in the location of activation to different face and body parts involved in biological movement. Taken from Allison et al. (2000), Figure 3. The area contained within the red circle approximates the region considered as pSTS.

]

A number of studies have identified a more anterior region of the STS in the processing of different gaze directions (Calder et al., 2007; Carlin, Calder, Kriegeskorte, Nili, & Rowe, 2011; Carlin, Rowe, Kriegeskorte, Thompson, & Calder, 2012). Allison, Puce and McCarthy (2000) suggest that these more anterior parts of the STS are activated by biological movement such as that of the hand and body. This was originally identified in macaques, in an anterior area of the STS which was shown to contain neurons which show responses to head orientation, gaze direction and body posture (Perrett et al. 1982; Perrett et al., 1985).

Using an adaptation paradigm, Calder et al. (2007) found that the right anterior STS (aSTS), after adapting to faces with a leftward gaze, showed increased responses to right compared with leftward gaze. The opposite pattern was seen after adapting to rightward gaze. In addition, work with monkeys and human patients with prosopagnosia identified deficits in the perception of gaze when the STS was lesioned (Campbell, Heywood, Cowey, Regard, & Landis, 1990).

In summary, the posterior STS is a key region involved in the processing of the changeable aspects of faces. It has been clearly implicated in the processing of a range of biological movements such as facial expressions and eye gaze. However the STS as a whole seems to be implicated in the processing of changeable aspects of faces, with the anterior STS being particularly associated with the processing of eye gaze.

1.3.5 The Amygdala

Haxby et al. (2000) identify the amygdala (Figure 1.5) as one of the regions in the extended system of their distributed neural system for face perception. They implicate the amygdala (along with the insula and other limbic system structures) in the processing of emotion from faces. From an evolutionary perspective, the amygdala is an old structure in the brain and this subcortical region is present in most vertebrates.

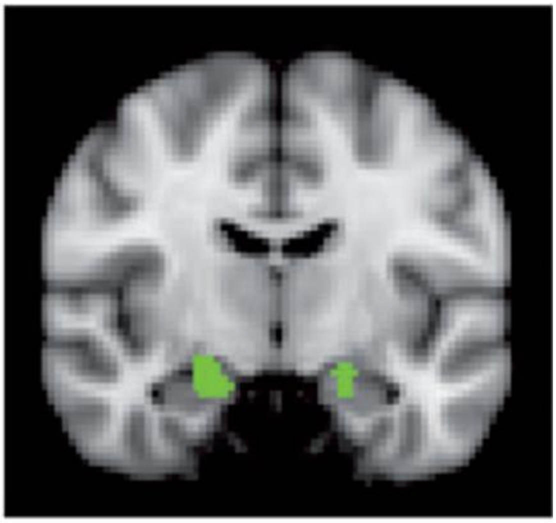


Figure 1.5. Face responsive regions in the left and right amygdala. Significantly active voxels were identified by a contrast of faces > non-faces (objects, places and Fourier scrambled faces. Adapted from Mattavelli et al. (2013), Figure 3.

Much of the previous discussion of facial expressions has focused on their relevance for social communication, however information gained from facial expressions can also be a key signal of threat in the environment. Correct interpretation of these signals can be crucial for survival. Fearful facial expressions suggest there may be a threat in the environment, and an angry facial expression may suggest that that individual is a potential threat to you.

The amygdala has been implicated in the processing of these biologically relevant signals, particularly that of fearful stimuli (Adolphs, Tranel, Damasio, & Damasio, 1995). There is a large body of neuroimaging research which demonstrates higher responses in the amygdala to threat related stimuli, e.g. fearful and angry facial expressions, compared to other facial expressions (Calder, Lawrence, & Young, 2001; Mattavelli et al., 2014; Morris et al., 1996; Whalen et al., 2001).

Studies of patients with amygdala lesions show difficulties in processing fear (Adolphs et al., 1999; Adolphs, Tranel, Damasio, & Damasio, 1994; Broks et al., 1998; Calder, 1996). A particularly noteworthy case was studied by Adolphs and colleagues with patient SM (Adolphs et al., 1995, 1994). SM has Urbach-Wiethe disease, which resulted in bilateral amygdala damage. SM is impaired in her recognition of fearful facial expressions and seems unable to experience fear despite being exposed to fearful stimuli and seemingly understanding that these would be classed as fearful to other people (Feinstein, Adolphs, Damasio, & Tranel, 2011). Despite the lack of fear recognition, SM is still able to experience other emotions.

However, it has been proposed that the amygdala is not required for the experience of fear. Feinstein et al. (2013) recently demonstrated that CO_2 inhalation triggered fear and panic attacks in SM and two other patients with bilateral amygdala damage. This suggests that it is perhaps the triggering mechanism that is not functioning correctly – they were able to trigger fear from internal influences, but it is the triggering from the external environment that is not functioning correctly.

Studies involving the processing of fear have demonstrated that the eye region is sufficient a cue in itself to achieve accurate recognition of the fearful facial expression. Neuroimaging work has shown the key role of the eye region in the processing of fearful faces in the amygdala (Morris, deBonis, & Dolan, 2002; Whalen, 2004) and research with patient SM demonstrated her impaired fear recognition was a result of a lack of fixation on the eyes when judging facial expressions (Adolphs et al., 2005). Interestingly, when SM was told to fixate on the eyes when completing the recognition task, her recognition of fear reached the same level as control subjects, adding weight to the importance of this region for fear recognition. However work by Asghar et al. (2008) demonstrated that although the eye region is sufficient for fear processing in the amygdala, responses to fearful faces with masked eyes showed that it is not the only cue involved.

There have also been a number of studies looking at how rigid movements of the eyes modulate the response of the amygdala. As discussed above, the amygdala is widely believed to be responsive to threat-related emotional expressions (e.g. fearful and angry facial expressions) and eye-gaze gives important information about where the fear or anger may be directed or what the source of the threat may be. Clearly, someone looking angrily at you is more of a threat than them looking angrily at something else. On the other hand, someone looking fearful at you is probably less of a concern than them looking fearfully at something you cannot currently see.

Adams and Kleck (2003) used pictures of faces expressing either fearful or angry faces, coupled with either direct or averted eye gaze. They found that angry faces with direct gaze and fearful faces with averted gaze were more quickly and accurately identified than angry faces with averted gaze and fearful faces with direct gaze. Thus supporting the assumption that aggression directed at the perceiver, or fear directed at the environment is more pertinent to attend to than aggression directed elsewhere and fear directed at the perceiver. Adams, Gordon, Baird, Ambady and Kleck (2003) used fMRI to address whether the amygdala's response could be modulated by direct and averted gaze in fearful and angry facial expressions. Adams et al. (2003) suggested the amygdala's response may be increased by threat-related ambiguity (fear with direct gaze, and anger with averted gaze) when compared to clear threat (fear with averted gaze, and anger with direct gaze). They found that there were no differences in the responses to fear and anger as a function of gaze in the right amygdala, however they did find significantly higher responses in the left amygdala to displays of ambiguous threat over clear threat.

In summary, the amygdala is implicated in the processing of facial expressions, particularly those which convey threat (fear and anger). Not only does the amygdala respond to these facial expressions, higher responses can be seen when the message conveyed by these expressions becomes ambiguous.

1.3.6 The Inferior Frontal Gyrus

The role of the IFG in face perception is less clear than the regions discussed above. Although not specifically implicated in the Haxby et al. (2000) distributed neural model for face processing, the IFG has been proposed to be part of the extended face processing network (Davies-Thompson & Andrews, 2012; Ishai, 2008). Ishai et al. (2008) propose a bidirectional connection between IFG and other core and extended regions implicated in face processing (Figure 1.6).

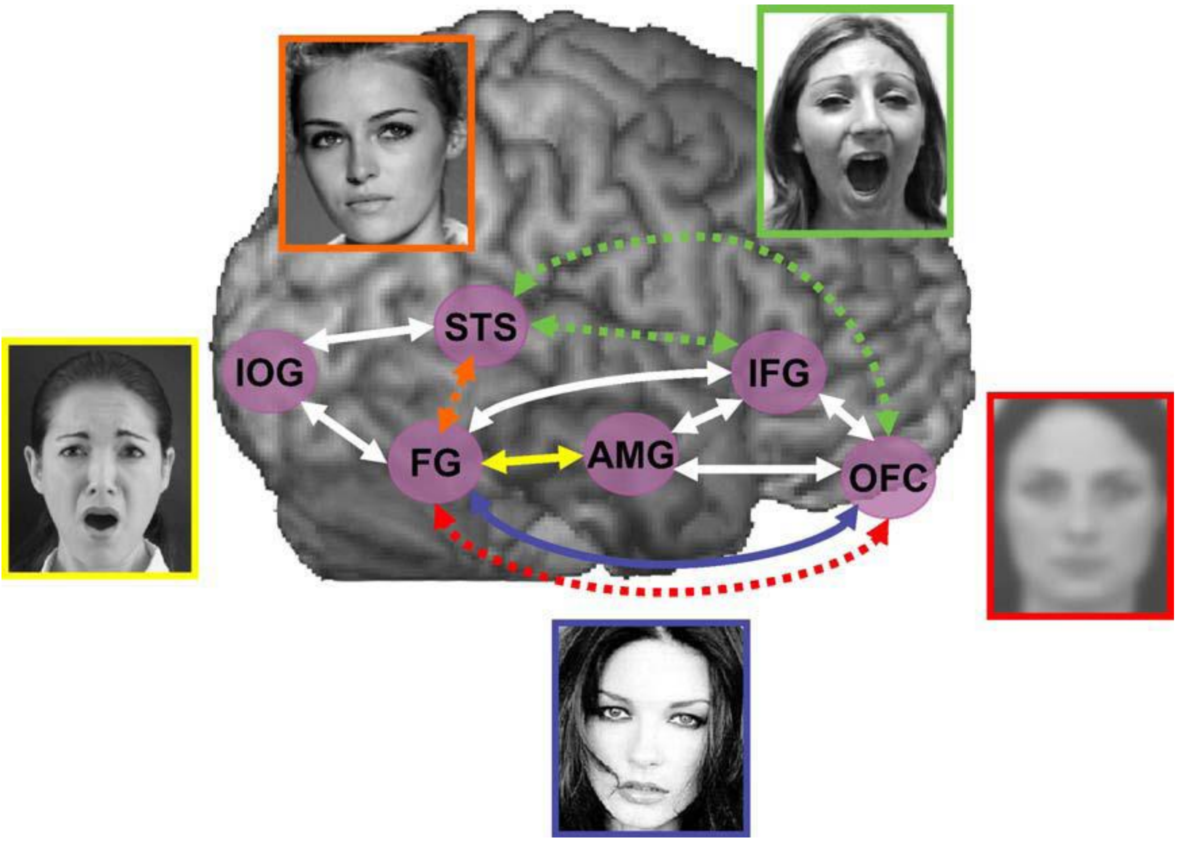


Figure 1.6. Model of neural coupling between core and extended regions as proposed by Ishai et al. (2008). Reciprocal connections are proposed between all regions, although not necessarily of equal strength. Viewing emotional faces increases the effective connectivity between the FG and the AMG (yellow), whereas viewing famous, attractive faces increases the coupling between the FG and the OFC (blue). New predictions are shown in dashed arrows: Attention to gaze direction would increase the coupling between the STS and the FG (orange); viewing animated faces would increase the coupling between the STS and the IFG/OFC (green); viewing indeterminate, low-spatial frequency faces would result in increased effective connectivity from the OFC to the FG (red). Taken from Ishai et al. (2008), Figure 2.

Inferior occipital gyri (IOG); fusiform gyrus (FG); superior temporal sulcus (STS); amygdala (AMG); inferior frontal gyrus (IFG); oribitofrontal cortex (OFC).

A study using dynamic casual modelling (DCM) to look at the neural coupling between the core and extended regions of the face processing network has demonstrated a strong feed-forward causal influence on the IFG from the fusiform gyrus (Fairhall & Ishai, 2007), supporting the neural coupling model proposed by Ishai et al. (2008).

A number of studies have implicated the IFG in the semantic aspects of face processing, particularly that of visual imagery (Ishai, Ungerleider, & Haxby, 2000; Ishai, 2002, 2008; Leveroni et al., 2000). In addition, the IFG has also been specifically implicated in the processing of facial expressions (Carr, Iacoboni, Dubeau, Mazziotta, & Lenzi, 2003; Dapretto et al., 2006; Ishai, Schmidt, & Boesiger, 2005). Dapretto et al., (2006) studied mirror neuron dysfunction in a group of children with autism spectrum disorders. The children with autism imitated emotional expressions as well as controls and there were no differences between the groups in the amount of time spent fixating the face and eyes. However there was a distinct difference in the neural responses. The children with autism showed no activity in the mirror area of the pars opercularis (a region of the IFG). This could not be explained by these children failing to attend to the face stimuli as they showed comparable activity in other face processing regions such as the fusiform gyrus and amygdala, to that of the control group. People with autism typically display deficits in understanding and interpreting emotional states in others and these results support the hypothesis that this could be a result of a dysfunctional mirror neuron system and clearly implicate the involvement of the IFG in the processing of facial expressions. However, recent studies suggest the difficulties in emotion processing in autism are the result of a co-occurring condition alexithymia, which is characterised by an impairment in identifying and describe one’s own emotion, which results in reduced empathy and ability to recognise emotion in others (Bird & Cook, 2013; Cook, Brewer, Shah, & Bird, 2013).

Although generally associated with the processing of facial expressions, the role of the IFG in the right and left hemisphere may be slightly different. A study by Ishai et al. (2005) found significantly higher response amplitudes in the right IFG to emotional faces over famous faces. However, the opposite was found for the left IFG. This is supported by a recent study by Flack et al. (2015) whereby responses to changes in facial expression were seen in the right IFG, but not the left.

To summarise, the IFG is becoming more widely thought of as part of the face processing system. Although fewer studies have focused on this region in comparison to other areas involved in face perception, it has been clearly implicated in the processing of facial expressions. The IFG may also have other roles involved in the processing of dynamic faces and has a feedback relationship with other regions in the network.

1.4 The neural representation of facial expression and viewpoint

Although much research has identified regions of the brain, such as the pSTS and amygdala, as being responsible for the processing of changeable aspects of faces, the question of how these aspects are represented within these regions is still unanswered.

Numerous studies have identified discrete regions specialised for processing different aspects of faces, yet it has been suggested that there may be a more distributed topographic organisation for the representation of these changeable aspects (Haxby et al. 2000). Furthermore, Haxby et al.'s (2000) model covers both rigid and non-rigid movements under the domain of the STS. These types of movement are clearly quite different, and it raises the question of how these different types of movement may be represented within the STS and beyond into the extended face processing system.

With respect to rigid and non-rigid movements, although studies have identified regions where neural responses to both of these types of movement occur, there has been little research to determine the extent to which these responses reflect distinct or overlapping representations. The work discussed previously, demonstrating the modulation of neural responses to non-rigid movements as a result of rigid movements suggests that there may be some degree of overlap (Adams & Kleck, 2003; Hadjikhani, Hoge, Snyder, & de Gelder, 2008). Indeed, the STS has been suggested to have distinct yet overlapping neural systems for the processing of facial expression and gaze direction (Engell & Haxby, 2007). However previous studies have not directly addressed the underlying neural patterns of response, and whether this pattern of response is distinct for each of these changeable aspects. It raises the question of whether there are distinct populations of neurons coding rigid and non-rigid movements, or whether there are overlapping fine-scale representations.

Multi-variate pattern analysis (MVPA) was first proposed by Haxby et al. (2001) and was used to demonstrate that neural patterns of response can be used to distinguish between object categories. They suggest that faces and objects have distinct patterns of neural response and these neurons are involved in representing many object categories. Haxby and colleagues (2001) study was founded on the basis that if a stimulus category evokes a distinct pattern of activity, then responses to multiple observations of that category should be more similar than responses to multiple observations of different categories. In essence, patterns of response to independent observations of a happy facial expression should be more similar to each other, than to independent observations of fearful, disgusted and sad facial expressions. Haxby and colleagues used correlation to measure the similarity in these patterns of response to object categories. The data were split into odd and even runs for each participant and the patterns of response to the object categories were compared both within and between category. Using this technique, they were able to demonstrate that patterns of response to a given category were relatively consistent within individuals and distinguishable from other categories. However, it would seem logical to suggest that the neural representation of objects might occur in a common way across individuals and not just within individuals, although it has also been proposed that these neural patterns occur in an idiosyncratic fashion (Haxby, 2012). The MVPA technique described in Chapter 2 and used in Chapters 4, 5 and 6, allowed this question to be addressed – whether patterns of activation might be unique to an individual or whether there exists a more common representation across individuals.

In addition to work looking at the representation of object categories, MVPA has also been used to identify patterns of response to aspects of faces such as eye gaze direction and dynamic head turns (Carlin, Calder, Kriegeskorte, Nili, & Rowe, 2011; Carlin, Rowe, Kriegeskorte, Thompson, & Calder, 2012) and has therefore been demonstrated as a reliable method for distinguishing aspects within object categories.

MVPA considers the response of multiple voxels simultaneously and as a result does not discard weaker responses that might be relevant, as seen in more traditional univariate analysis. For example, small but significant responses can be seen to non-face objects, in a 'face area', but are often not considered further (Haxby, 2012). This approach has led to an idea of modularity whereby each object category may have a dedicated region or module for processing e.g. the FFA for faces (Kanwisher et al., 1997). Given the large number of objects in our environment, and the number of distinctions that can be made within an object category, it is likely more plausible that these could be represented by overlapping fine-scale representations rather than peak responses in category-specific regions.

In summary, MVPA has led to a reconsideration of how different categories of objects might be represented. MVPA considers the overall neural pattern of response, and therefore provides a promising route to determine how multiple object categories and aspects within these object categories, could be represented across a limited area of cortex.

1.5 Thesis aims

One of the clear next steps is to probe further the underlying neural coding of the changeable aspects of faces, in order to better understand how face responsive regions of the brain process this important information. As discussed above, MVPA allows the opportunity to explore this coding, and gives an alternative way of looking at this information. This thesis will use both MVPA and complementary univariate techniques to address the representation of expression and viewpoint. Much of the research discussed above, focused on facial expression and eye gaze, however changes in eye-gaze clearly demonstrate a much more subtle change in the face (and an image as a whole) than facial viewpoint. Despite this, eye gaze conveys the same kind of information to the perceiver as changes in viewpoint, suggesting there may be a substantial overlap in how face responsive regions of the brain process this information.

As discussed above, viewpoint and expression play crucial roles in face perception. This thesis aims to take a novel approach to exploring the underlying organising principles of these changeable aspects of faces. The main aims of this thesis are 1) to investigate the neural representation of facial expression in the core and extended regions of the face processing network as identified by Haxby et al. (2000). 2) to investigate the neural representation of facial viewpoint in those same regions, 3) to compare the representation of expression and viewpoint and assess whether these changes have distinct representations. This thesis addresses significant questions about the organising principles and processing of faces in the human brain. The aims specifically addressed by the following experimental chapters are summarised below.

Chapter 2 – this chapter provides a brief overview of the stimuli and general methods used throughout the following experimental chapters.

Chapter 3 – the first experimental chapter uses fMR adaptation to address how facial expression is represented in core and extended regions of the face perception network. More specifically, this chapter asks whether these regions represent facial expression in a holistic or feature-based way. Previous behavioural research has demonstrated that facial expressions are processed holistically. Whilst the neural underpinnings of the holistic processing of identity have been explored, this has not been addressed for facial expressions.

Chapter 4 – this chapter investigates the representation of changes in expression and viewpoint across the core and extended regions. Although there has been much research into the neural processing of facial expressions and viewpoints, the underlying representation of these distinct changes is still unclear. Using multivariate pattern analysis, this chapter asks whether we have distinct representations of changes in expression and viewpoint. In addition, an adaptation experiment addresses the relative representation of expression and viewpoint in the key regions.

Chapter 5 – this chapter addresses how the brain represents individual viewpoint directions. This chapter addresses a theory proposing that representation moves from a viewpoint-specific representation in earlier visual areas, to partial viewpoint invariance, through to full viewpoint invariant representation in more anterior regions. Evidence to support this theory would provide a significant contribution to the understanding of how we develop invariant representations of faces.

Chapter 6 – the final experimental chapter seeks to address the neural representation of specific combinations of expression and viewpoint. This chapter uses behavioural and neural data to ask how viewpoint modulates the response to expression and vice versa. Whilst there is some support for distinct decodable patterns of response to facial expressions and viewpoint directions, whether this is possible for specific combinations of expressions and viewpoints (e.g. a right facing fearful face) has not been addressed.

Chapter 7 – the final chapter of this thesis gives a general discussion of the findings of the previous experimental chapters, and how this influences our understanding of the representation of faces in the human brain.

Chapter 2 – General Methods

2.1 Stimuli

Careful consideration was given to the choice of face stimuli used throughout the experiments described in this thesis. Great care must be taken in any experiment to ensure the stimuli are appropriately controlled. Two face databases were chosen, the FEEST set, used in Chapter 3, and the Radboud Database, used in Chapters 4 - 6.

2.1.1 FEEST Set

In Chapter 3, faces from the Young et al. (2002) FEEST set, derived from the Ekman and Friesen (1976) Pictures of Facial Affect (POFA) were used. The POFA were originally developed to use in studies looking at facial expression recognition across cultures, (described in Chapter 1, section 1.2.3. *Facial Expressions*). Ekman and Friesen ensured that actors posing the expressions did so in a consistent way, by using specific facial muscles to pose the expressions. This system was later published as The Facial Action Coding System (FACS), (Ekman & Friesen, 1978) as a system for describing and distinguishing between different facial movements. The value of this system means that facial expressions can be described on a muscle by muscle basis, meaning it is possible to ensure consistency within expressions across individual models.

Images in the Young et al. (2002) FEEST set were taken from Ekman and Friesen's POFA. Young et al. (2002) selected images from the POFA to ensure that all actors chosen for the FEEST set (10 actors) had images of all six expressions (fear, anger, happiness, disgust, sadness and surprise) plus a neutral pose. For the experiments in Chapter 3, two individuals posing four facial expressions (fear, anger, happiness and disgust) were used to create the stimuli. The individuals used in Chapter 3 were selected on the basis of a high recognition rate for all expressions and consistency of the action units used to pose each expression (Young, Perrett, Calder et al. 2002).

2.1.2 Radboud Database

For Chapters 4-6 it was necessary to use stimuli where the faces changed in viewpoint direction. All the faces used in the FEEST set are forward facing and so a different set of well controlled face images were required. Chapters 4-6 used images from the Radboud Database which has a large set of faces which vary in expression, viewpoint and eye gaze direction (Langner et al., 2010). These images were also FACS rated and the models were taught how to pose the facial expressions using the standardised muscles as defined by FACS. The expressions are therefore reasonably consistent across models. They also used a simultaneous camera capture which meant they could take photos of the various poses from five different viewpoints (Figure 2.1). This means the other aspects of the face e.g. expression, are all consistent across the five viewpoints, which was an important factor when addressing the modulation of expression perception by viewpoint (Chapter 6).

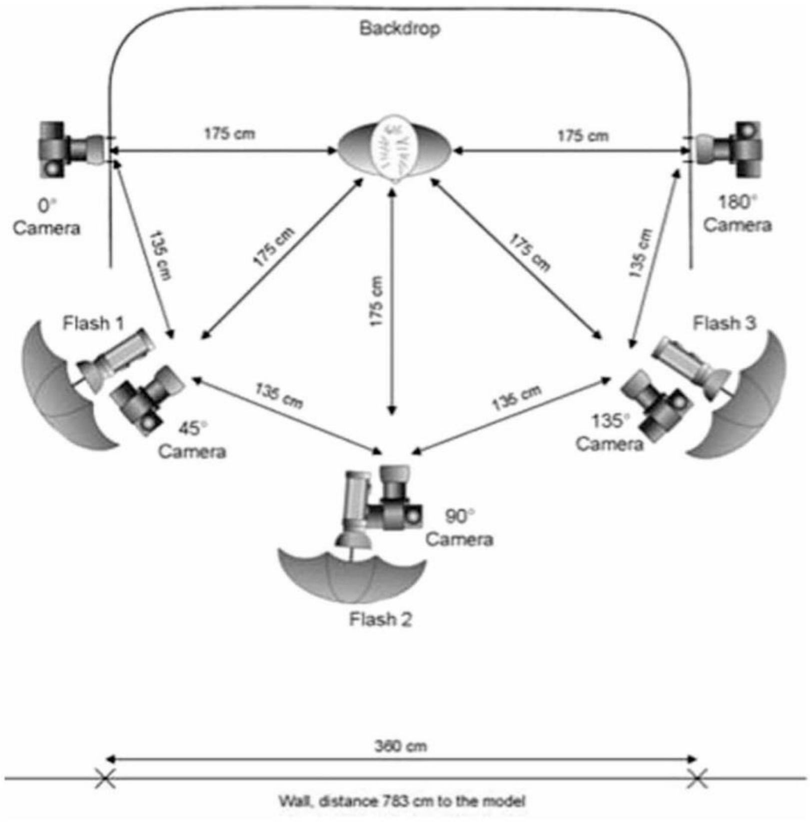


Figure 2.1. Technical setup of the photo shoot for the Radboud Database (reproduced from Langer et al. 2010, Figure 3).

2.2 Region of Interest Definition

Defining regions of interest (ROIs) is a crucial step in any fMRI analysis. For the results of any analysis to be meaningful, it is important to define the boundary of any region you are interested in to a high degree of certainty. There are many ways to do this but they are broadly divided into two types, structurally defined and functionally defined. Structurally defined ROIs can be defined using the individual's anatomical architecture, or defined at the group level, for example with an anatomical atlas which has been created by calculating the average location of particular structures, across a number of individuals. An example of this is the Harvard Oxford Atlas which is a probabilistic atlas whereby each structure is represented as a standard image with values from 0 to 100, according to the cross-population probability of a particular voxel being in that structure. One of the main advantages of using structurally defined ROIs is that it avoids assumptions about the function of a particular structure. However, areas of interest often have functional subdivisions which are more difficult to define from the anatomical architecture and so functionally defined ROIs are more useful in this case. This is particularly relevant for face perception research where the main face-responsive regions of the OFA, FFA and pSTS correspond to areas of the posterior occipital cortex, inferior fusiform gyrus and superior temporal lobe. The ROIs used in this thesis are defined functionally and the different methods used in Chapter 3 and Chapters 4-6, are described below.

2.2.1 Functional localisation at the individual level

In Chapter 3, ROIs were defined for each individual, and the subsequent analysis was performed within each individual's own functional ROIs. The advantage of functional localisation on an individual basis is that it ensures with a relative degree of certainty that you are targeting the appropriate brain area for each person, therefore accounting for individual differences in the position of these areas. The initial stage requires a functional localiser which was conducted as part of the scan session for each participant.

As the aim of this thesis is to understand areas of the brain involved in face perception, the first stage was to use a functional localiser scan designed to produce activity in face-responsive regions of the brain – in particular the OFA, FFA and STS. The localiser had 3 stimulus conditions: faces, places, and Fourier phase-scrambled faces. The localiser scan had a block design, and each block lasted 9 seconds and contained 9 images from one of the localiser conditions. Each image was presented for 900 milliseconds with a 100 ms inter-stimulus interval (ISI). Each condition was repeated 4 times. The localiser scan was presented using Neurobehavioural Systems Presentation 16.3.

The face-selective regions were defined in each individual from the functional localiser data by using the average of the faces > places and faces > scrambled faces contrasts. The combined statistical maps were thresholded at p < .01 (uncorrected). For each individual, the OFA, FFA and pSTS were identified by contiguous clusters of voxels activated above threshold from the above contrast in posterior occipital cortex, inferior fusiform gyrus and superior temporal lobe. These functional areas were then used for all subsequent analysis in Chapter 3.

2.2.2 Functional localisation at the group level

In Chapters 4 – 6, a different technique was used to identify face-responsive ROIs. The functional localiser used was developed initially for the definition of functional ROIs in different fMRI studies. Because these studies used the same localiser, a unique opportunity arose to generate ROIs based on a large group of participants (n=83). These ROIs could then be used in different studies to explore the neural basis of face processing in corresponding regions of the brain.

The localiser had 5 stimulus conditions, with conditions 1-4 being of faces: (1) *same-identity, different-expression* (sIdE), (2) *different-identity, different-expression* (dIdE), (3) *same-identity, different-viewpoint* (sIdV), (4) *different-identity different-viewpoint* (dIdV) and (5) images taken from each face category, phase-scrambled in the Fourier domain. These face images were from the Radboud Faces Database, as described above (1.1.2. *Radboud Database*). The localiser scan had a block design, with each block lasting 6 seconds, and contained 5 images. Each image was presented for 1000ms with a 200ms black screen ISI. Each condition was repeated 5 times.

The face-responsive regions were defined at the group level from the localiser data by using an average of each face condition > scrambled condition. The combined statistical maps were thresholded at Z > 3.1, p < .001. Regions corresponding to the OFA, FFA and pSTS were defined based on contiguous clusters of significantly active voxels in the posterior occipital cortex, inferior fusiform gyrus and superior temporal lobe respectively. These ROIs were defined in each hemisphere, and then combined for the purpose of analysis. These functional ROIs were used in Chapters 4-6 with full methodological details being reported in Chapter 4.

2.2.3 Choice of ROI definition

As described above, different techniques were used to define the ROIs used in this thesis. The ROIs used in Chapter 3 were defined at the individual level, and Chapters 4-6 were defined at the group level. There are advantages to both methods, e.g. the relative certainty of targeting the correct functional area for a specific individual when defining at the individual level, and the relative convenience and high power gained from defining group level masks from a large sample size. For Chapter 3, it was more appropriate to use individual level masks to account for individual differences in the adaptation timecourses. However as discussed in the above section, a unique opportunity arose to create group level masks from a large dataset. The large number of participants (n=83) gave this analysis a high degree of power in identifying regions associated with face processing, and the use of these ROIs across different MVPA studies, allows direct comparisons to be made.

2.3 fMRI Analysis

2.3.1 Univariate Analysis

Traditionally fMRI analyses use a univariate general linear model (GLM) method, which is sensitive to local changes in the amplitude of response. GLM generates a model and measures the fit of this model to the data. The model is typically based on the timing of the stimuli presentation, and so a good fit between the model and the data suggests the variation in the data is caused by the stimuli presentation. This can be done for each experimental condition. Typically a boxcar model is used to fit the data, whereby an above zero response is expected when the stimulus is present, and a zero response is predicted when the stimulus is not present. This model is then convolved with the hemodynamic response function. This produces a regressor which is then regressed against the fMRI signal at the level of a single voxel. A statistical parametric map representing the regression coefficient, can then be created that indicates the degree of fit from the model for each voxel. This can be done for each experimental condition and the significance of this model fit is indicated by testing the regression coefficient against baseline, giving a p-value or z-value. These statistical parametric maps generated for an individual participant, can also be entered into a higher level analysis giving an estimate of model fit across a group of participants. Often, experimental conditions are contrasted with one another, to give a statistical parametric map which indicates differences in response between the two conditions.

2.3.1.1 fMR Adaptation

fMR adaptation is a widely used technique and works on the principal that if you show the brain the same stimuli repeatedly, the neurons which are responsive to that type of stimuli will become habituated. What this technique allows us to do is address whether for example in one voxel we have two populations of neurons coding different stimuli types, or whether we have one population of neurons coding both stimuli types. Figure 2.2 below illustrates an example of a standard fMRI experiment versus an adaptation experiment.

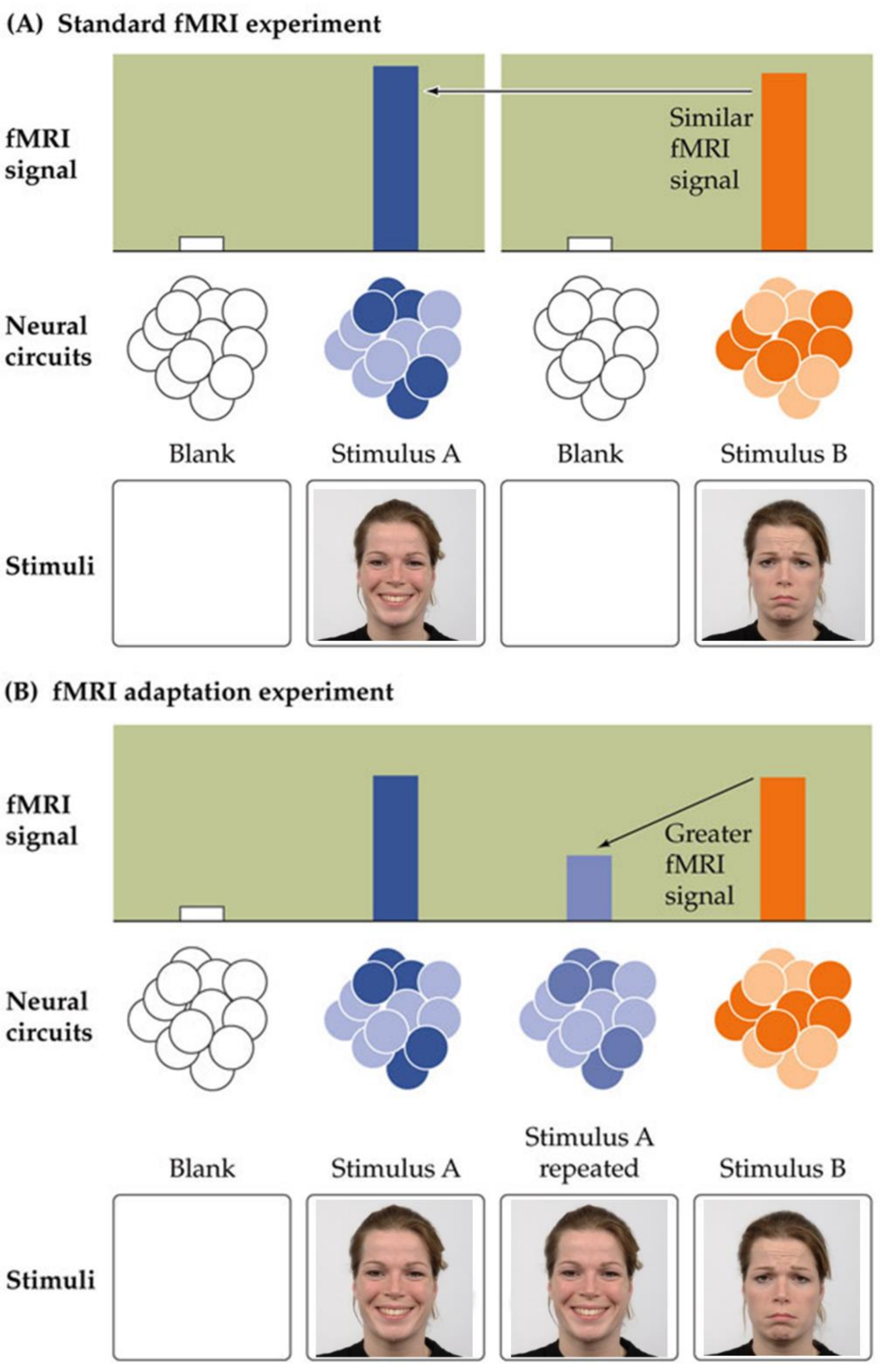


Figure 2.2. Examples of a standard fMRI experiment (A) versus an fMR adaptation experiment (B). Adapted from Principles in Cognitive Neuroscience, Box 15B. Purves et al. (2013).

In a standard fMRI experiment (Figure 2.2A) we could have a voxel which contains a population of neurons which is selective for the happiness facial expression. If a participant is shown this image, then you will get a response from this voxel of a particular magnitude. Within this voxel there is also a population of neurons which are selective for the sadness facial expression. If the participant is then shown an image depicting sadness, then this population of neurons could give a response of a similar nature to that of the previous expression. However in an adaptation paradigm, the presentation of a particular stimulus is repeated. So in an fMR adaptation paradigm (Figure 2.2B) the happiness expression is presented repeatedly and the response to this stimuli reduces. When the participant is then shown the sad face, as there is a separate population of neurons which are selective for sadness, a higher response for this stimulus is seen (these neurons were not habituated). Therefore, using an adaptation paradigm means the underlying neural coding for (in this example) the processing of facial expressions, can be more directly addressed. This method is used in Chapter 4, Experiment 2, to investigate the selectivity of response to facial expression and viewpoint across face-responsive regions of interest.

2.3.2 Multivariate Analysis

In addition to the univariate analysis traditionally used, multivariate analyses are now widely used in fMRI research. Multivariate techniques allow the analysis of patterns of response across multiple voxels simultaneously. Multi-voxel or multi-variate pattern analysis (MVPA) is typically performed on the parameter estimates that are output from univariate analysis. This is the method used for the MVPA analysis undertaken in Chapters 4-6.

2.3.2.1 MVPA Methods

As discussed in the previous chapter, MVPA was first proposed by Haxby et al. (2001) and was used to demonstrate that neural patterns of response can be used to distinguish between object categories. Haxby and colleagues used a correlation based method of MVPA, and although a number of other types of MVPA methods have developed since (classification algorithms, searchlight analysis), the correlation based method will be the focus of this thesis.

In correlation based MVPA, parameter estimates for each stimulus condition are used. The data is then split, and in the case of Haxby et al. (2001), this was done by splitting the data into odd and even runs for each participant. The data is then normalised to reduce the shared variance across the conditions. For example, if the stimulus conditions were happy faces, sad faces and fearful faces, the parameter estimates of these conditions are all likely to contain variance in the data which explains 'face'. By normalising the parameter estimates, this shared 'face' representation is removed (along with variance in the data which relates to more generic aspects e.g. a visual stimulus), leaving behind the pattern of response that is unique to the experimental condition. The normalisation process used in Chapters 4-6 involves subtracting the mean response across all experimental conditions, from each individual condition on a voxel by voxel basis. The normalised patterns of response can then be compared using pairwise correlations, both within and between category. An example taken from Haxby et al. (2001) can be seen in Figure 2.3 below.

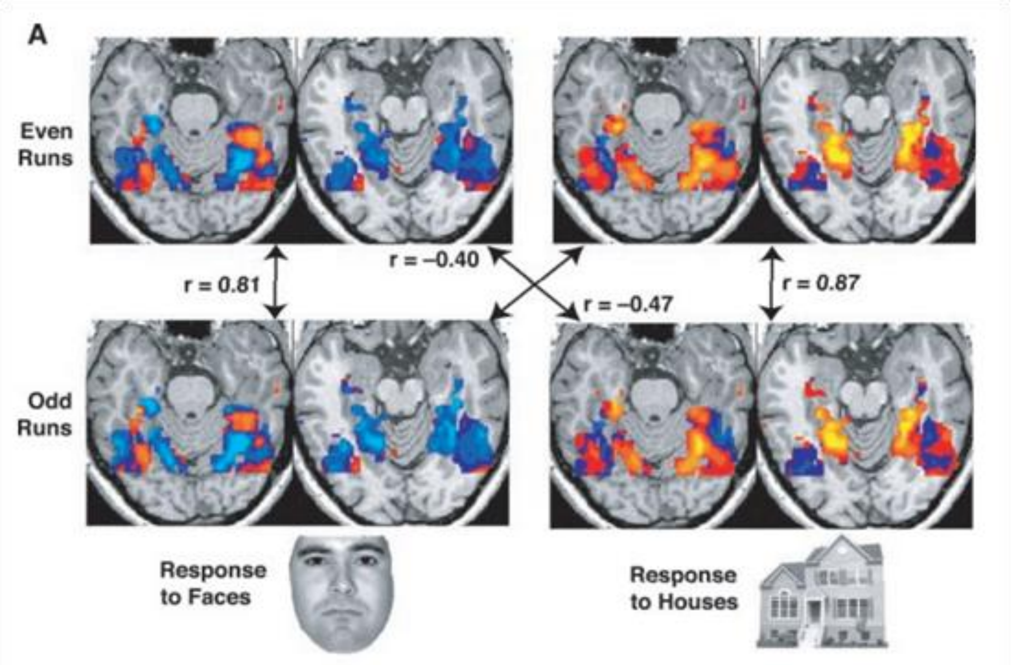


Figure 2.3. Normalised patterns of response to two stimulus categories (faces and houses) in the ventral temporal cortex. Pairwise correlations between the patterns of response to the two stimulus categories show higher correlations within-category (e.g. even run faces versus odd run faces), than between category (e.g. even run faces versus odd run houses). Adapted from Haxby et al. (2001), Figure 3.

If there is a reliable pattern of response to a given category, then higher correlations will be seen within category than between category. It this is the case it suggests that there is a high degree of similarity of the neural patterns within a given category which allows you to distinguish it from other categories.

2.3.2.2 MVPA LOPO

The MVPA method used throughout this thesis uses a leave-one-participant-out (LOPO) cross validation paradigm. The LOPO paradigm has the advantage of being able to determine the consistency of responses across individuals. The methodological details of the LOPO paradigm used in Chapters 4 – 6 are described below.

2.3.2.3 LOPO Methods

The pattern analyses used in Chapters 4 – 6 were performed using the PyMVPA toolbox (http://www.pymvpa.org/; Hanke et al., 2009). The parameter estimates from the univariate analysis were first normalised by subtracting the average response across all face conditions (the number of face conditions varied depending on the experiment - see the relevant chapter for details). The reliability of the neural patterns of response was then determined using a modified form of the correlation-based MVPA method devised by Haxby and colleagues (2001). The Haxby et al. method splits the data from each participant into odd and even stimulus runs and then correlates patterns of response to different conditions across each participant's odd and even runs. This procedure determines whether responses are consistent at the individual participant level. To determine consistency across participants, a modified version of this cross-validation paradigm was used whereby the patterns of response from each participant were compared to the patterns resulting from the group analysis with that participant left out. This LOPO method makes it possible to determine the consistency of the patterns of response across participants by measuring how similar each participant's responses were to those for the rest of the group (Rice, Watson, Hartley, & Andrews, 2014).

The parameter estimates for each experimental condition in each voxel were normalized by subtracting the mean response in each individual voxel across all relevant experimental conditions (in the case of Chapters 4-6, the face conditions). The group pattern is derived by entering all but one of the participants’ data into a higher-level group analysis (mixed effects, FLAME http://www.fmrib.ox.ac.uk/fsl). This group pattern of response for each condition is then correlated with the pattern from the participant who was omitted from the group. For each unique pair of conditions, the LOPO method is repeated *n* times (*n* being the number of participants), with a different participant being omitted from the rest of the group each time. In Chapters 4-6, paired samples t-tests were then used to test the difference between the within-condition and between-condition correlations. If a particular stimulus category evokes a distinct pattern of activity, then the within-condition correlations for the individual participant and rest of the group should be higher than the between-condition correlations. An example of the LOPO method can be seen in Figure 2.4.

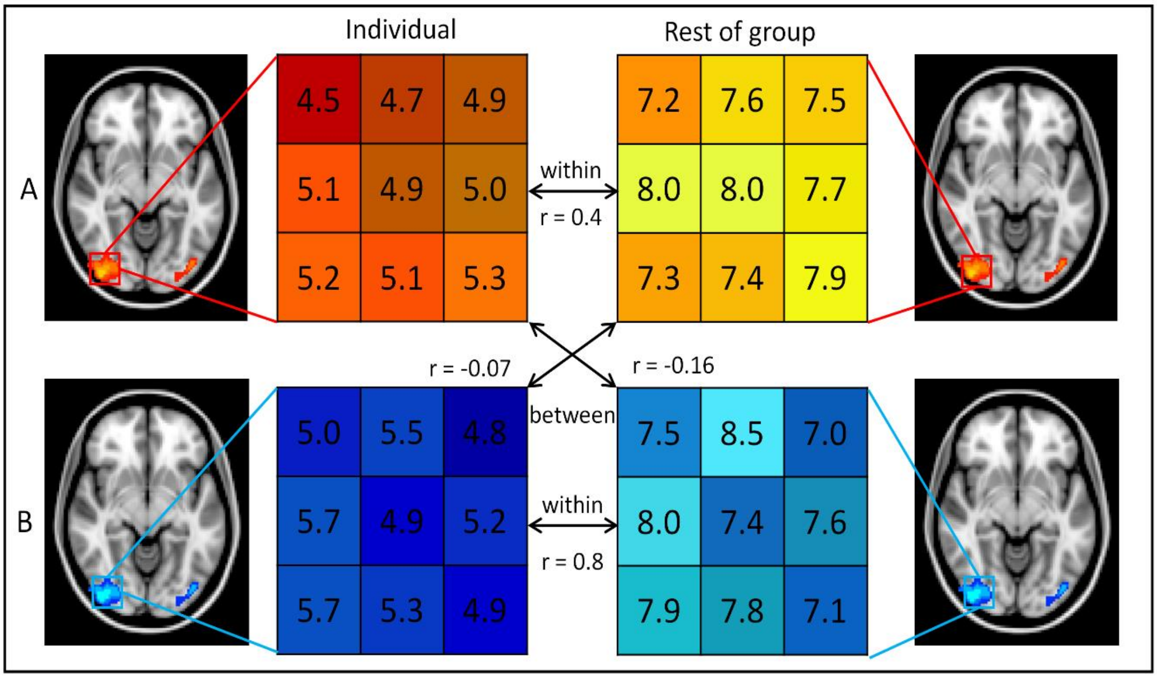


Figure 2.4. An example of the LOPO paradigm. Patterns of neural response are taken from two stimulus conditions, A and B. These patterns of response are restricted to a given region of interest. The patterns are then compared using pairwise correlations across the splits in the data (individual data versus group data) within category (e.g. condition A in the individual versus condition A in the group), and between category (e.g. condition A in the group versus condition B in the group). Higher within-category correlations versus between category correlations, indicates a level of similarity in the neural patterns within a given category which allows you to distinguish it from other categories.

2.3.2.4 Respresentational Similarity Analysis (RSA)

Representational similarity analysis has recently become a popular way in addressing the underlying functional representations of patterns of neural response generated from MVPA. RSA allows you to compare the neural representation to various models of representation, for example computational models or more basic models aimed at representing behaviour (Kriegeskorte, Mur, & Bandettini, 2008). This is done in the form of similarity matrices, whereby two matrices can be compared to one another. The MVPA correlation matrix is compared to the model matrix, by correlating each of the elements between the two matrices. If the two matrices are similar to one another, then this suggests the model is able to predict the similarity between the conditions in the MVPA matrix. This model can then be considered a good predictor of the underlying functional representation of the neural data. In this thesis, models are generated in various ways. For example in Chapter 4, binary models are created based on predictions about underlying neural representation. In Chapter 6, models were created based on behavioural perceptual ratings of the similarity of facial expressions and facial viewpoints.

2.4 Methods Overview

In addition to behavioural paradigms that will be described in the relevant chapters, a combination of the methods described above are used throughout this thesis. Chapters 3 and 4 both use univariate methods, including the fMR adaptation paradigm. Chapter 3 uses fMR adaptation to address how facial expressions are represented in face selective areas and Chapter 4 uses fMR adaptation to compare the relative representation of facial expression and facial viewpoint in the face-responsive ROIs. MVPA LOPO is used in Chapters 4 – 6, to further explore the representation of expression and viewpoint, looking at whether these changeable aspects have distinct or overlapping neural representations.

Chapter 3 – Holistic and Feature-Based Responses to Facial Expression

This chapter is adapted from: Flack, T. R., Andrews, T. J., Hymers, M., Al-Mosaiwi, M., Marsden, S. P., Strachan, J. W. A., Trakulpipat, C., Wang, L., Wu, T., & Young, A. W. (2015). Responses in the right posterior superior temporal sulcus show a feature-based response to facial expression. *Cortex, 69,* 14-23.

3.1 Abstract

Chapter 3 aims to address how facial expression is represented in face-responsive regions of the human brain. In particular, the face-responsive region of the right posterior superior temporal sulcus (pSTS) plays an important role in analysing facial expressions, however it is less clear *how* facial expressions are represented in this region. In this chapter, the face composite effect was used to explore whether the pSTS contains a holistic or feature-based representation of facial expression. Images were created from the top and bottom halves of faces posing different expressions. These images either had the top and bottom halves aligned into a face-shape, or misaligned. In Experiment 1, participants performed a behavioural matching task in which they judged whether the top half of two images was the same or different. The ability to discriminate the top half of the face was affected by changes in the bottom half of the face when the images were aligned, but not when they were misaligned. This shows a holistic behavioural response to expression. In Experiment 2, fMR-adaptation was used to ask whether the pSTS has a corresponding holistic neural representation of expression. Aligned or misaligned images were presented in blocks that involved repeating the same image or in which the top or bottom half of the images changed. Increased neural responses were found in the right pSTS regardless of whether the change occurred in the top or bottom of the image, showing that changes in expression were detected across all parts of the face. However, in contrast to the behavioural data, the pattern did not differ between aligned and misaligned stimuli. This suggests that the

pSTS does not encode facial expressions holistically. In contrast to the pSTS, a holistic pattern of response to facial expression was found in the right inferior frontal gyrus. Together, these results suggest that pSTS reflects an early stage in the processing of facial expression in which facial features are represented independently.

3.2 Introduction

Interpreting the facial expressions of others is important to effective social interaction (Bruce & Young, 2012). Facial expressions result from characteristic patterns of movement of the facial muscles that can easily be seen in static photographs (usually showing the apex of the movement itself) or in videos (Johnston et al., 2013). However, little is known about how expressions are encoded at the neural level. The most widely-used neural model of face perception (Haxby et al., 2000) proposes that the superior temporal sulcus (STS) is a key neural structure in the perceptual analysis of facial expressions, and this is borne out by a number of studies that have implicated STS in neural responses to expression (Calder & Young, 2005; Psalta, Young, Thompson & Andrews, 2014) and social perception from visual cues (Allison, Puce & McCarthy, 2000).

Relatively few studies address the question of *how* STS encodes expression. Said et al. (2010) were able to demonstrate that patterns of activation to different facial expressions across voxels in posterior STS (pSTS) were correlated with the rated perceptual similarities of the expressions themselves, suggesting that the functional organisation of pSTS reflects this underlying perceptual structure. Similarly, Harris et al. (2012) found that right pSTS responded to changes in facial expression regardless of whether or not these changes crossed or remained within emotional category boundaries, which again suggests a form of encoding that is largely driven by the perceptual input. Importantly, Harris et al. (2014) showed that right pSTS is relatively insensitive to contrast reversal, which implies that the critical perceptual input for pSTS involves feature shapes. Contrast reversal is known to have a dramatic effect on face identity recognition, but it has relatively little effect on the recognition of expression because information about feature shapes that is critical to interpreting facial expressions is conveyed through the position of edges that remain largely invariant to contrast reversal (Bruce & Young, 1998).

Here, the study of the perceptual representation used by pSTS is taken a step further by asking whether it represents features such as the eyes and mouth independently from each other, or as part of a perceptual whole (the face). The critical test of holistic processing that is used for this purpose is the expression composite effect. Composite effects have been demonstrated in many studies of facial identity perception (Young, Hellawell & Hay, 1987; Rossion, 2013), but their extension to understanding facial expression perception is less well-known. The paradigm involves combining the top half of one facial expression with the bottom half of another expression and determining whether this combination of different parts results in the perception of a new whole expression (Calder & Jansen, 2005; Calder et al., 2000; Palermo et al., 2011; Prazak & Burgund, 2014). The critical test of holistic perception involves contrasting performance between images in which the top and bottom halves are aligned into a highly face-like overall configuration, or misaligned so that they are less face-like. Contrasting aligned and misaligned versions of composite images created from the top and bottom parts of different facial expressions makes it possible to differentiate responses based on face features, which will be equivalent across aligned and misaligned image variants, from holistic responses that will only be evident for aligned and not for misaligned images.

In this chapter, the facial expression composite effect was used to investigate whether neural responses to facial expression in right pSTS reflect feature changes or are dependent on the face as a perceptual whole. To do this, it was important to first establish in a behavioural study, that the stimuli and presentation parameters that were intended to be used in fMRI, elicited a robust expression composite effect. Neural responses in right pSTS to composite expressions in which the top (eye region) and bottom (mouth region) parts were aligned into an overall face-like configuration were compared with neural responses to misaligned stimuli created by shifting one part horizontally with respect to the other (see Figure 3.1). Misalignment still allows the separated parts of the face to be encoded as features, but it interferes with the integration of expressive information from the eye and mouth region into a perceptual whole (Calder et al., 2000).

This fMRI experiment used a block design adaptation paradigm in which participants viewed blocks comprising a series of facial expressions that were all the same (no change condition) or that varied across the top half of each image (top change condition) or across the bottom half of each image (bottom change condition). During these blocks, participants were asked to fixate between the eyes (i.e. in the top half of each face) and further to encourage fixation they had to detect the presentation of an occasional small red spot at the fixation point. The no change condition, with identical stimuli throughout the block, served as a baseline that will lead to maximal adaptation of neural responses, and the top change or bottom change conditions measured any release from adaptation in neural regions that can encode these changes. The stimuli were aligned into overall face-like composites, or horizontally misaligned so that they were not face-like (see Figure 3.1), allowing to establish whether the pattern of neural responses across conditions involving no change, top change, or bottom change was dependent on the presence of a face-like (aligned) configuration. The prediction was that the pSTS would respond holistically to facial expressions. This would be shown by a higher response to the bottom change condition compared to the no change condition, but critically this would only be when the stimuli were aligned into a face-like configuration.

3.3 Materials and Methods

3.3.1 Participants

Sixteen participants took part in Experiment 1 (8 male, 8 female, mean age 27.6 ± 4.4). Twenty-seven participants took part in Experiment 2 (17 male, 10 female, mean age 24.7 ± 5.0). All participants had normal or corrected-to-normal vision, with no known history of neurological disorder and no abnormalities that were immediately evident from structural MRI in Experiment 2. Written consent was obtained from all participants and the studies were approved by the York Neuroimaging Centre Research Ethics Committee and the Department of Psychology Ethics Committee at the University of York. One participant was removed from the fMRI analysis due to excessive head movement.

3.3.2 Experiment 1

3.3.2.1 Stimuli and Design

Experiment 1 was an initial behavioural study used to validate key procedural parameters which would then be used in Experiment 2. It was important to conduct this experiment to ensure the expression composite effect could be demonstrated with the chosen parameters. Stimuli consisted of aligned composite and misaligned non-composite images of greyscale faces which either had; 1) the same top and bottom half (no change); 2) the same bottom half with the top half varying in expression (top change); 3) the same top half with the bottom half varying in expression (bottom change). The combination of aligned and misaligned versions of these 3 conditions led to 6 conditions in total. Each of these 6 conditions involved 24 trials.

The top and bottom half images were separated by a gap of 5 pixels, in line with the procedural strictures of Rossion (2013). Top and bottom half face images were derived from Ekman faces taken from the FEEST set (Young et al. 2002). Two individuals posing four facial expressions (fear, anger, happiness and disgust) were used to create the stimuli. These individuals were selected on the basis of a high recognition rate for all expressions and consistency of the action units used to pose each expression (Young et al. 2002).

Aligned or misaligned images were presented in sequential pairs in which both members of the pair had aligned constituent parts or both had misaligned parts. In misaligned pairs the offset was to the left in half the trials, or to the right in the other half. Images were presented using an LCD monitor, approximately 57 cm from the participant. Aligned images were approximately 5° x 8°, and misaligned images were approximately 8° x 8°. The images were presented for 750 ms each, with a 750 ms inter-stimulus interval. The two images in each sequential pair were always made from parts of the same individual's face, so that face identity was not a confound in the experiment, but the top or bottom parts could differ in expression. Images for the behavioural experiment were presented using PsychoPy2 (Peirce, 2007).

Participants were instructed to only look at the top half of the face. There was a fixation cross located between the eyes on each ISI and a chin rest was used to help participants maintain fixation on the top half of the images. Participants had to judge whether the top half of the image was the same (identical) or different (in any way) across the pairs of images. Participants could respond as soon as the second image appeared, and were given a maximum of 3 seconds to respond.

3.3.3 Experiment 2

3.3.3.1 Stimuli and Design

Experiment 2 used a block design fMR-adaptation paradigm. In order to identify face-selective regions for each individual, a localiser scan was conducted prior to the experimental scan. The localiser had 3 stimulus conditions: faces, places, and Fourier phase-scrambled faces. Each localiser scan block lasted 9 seconds and contained 9 images from one of the localiser conditions, with each image being presented for 900 milliseconds and a 100 ms inter-stimulus interval (ISI). Each condition was repeated 4 times. Images used in the localiser scan were presented using Neurobehavioural Systems Presentation 16.3.

For the main fMR-adaptation scan, Experiment 2 had the same 6 stimulus conditions as Experiment 1, this time presented in a block design (Figure 3.1).

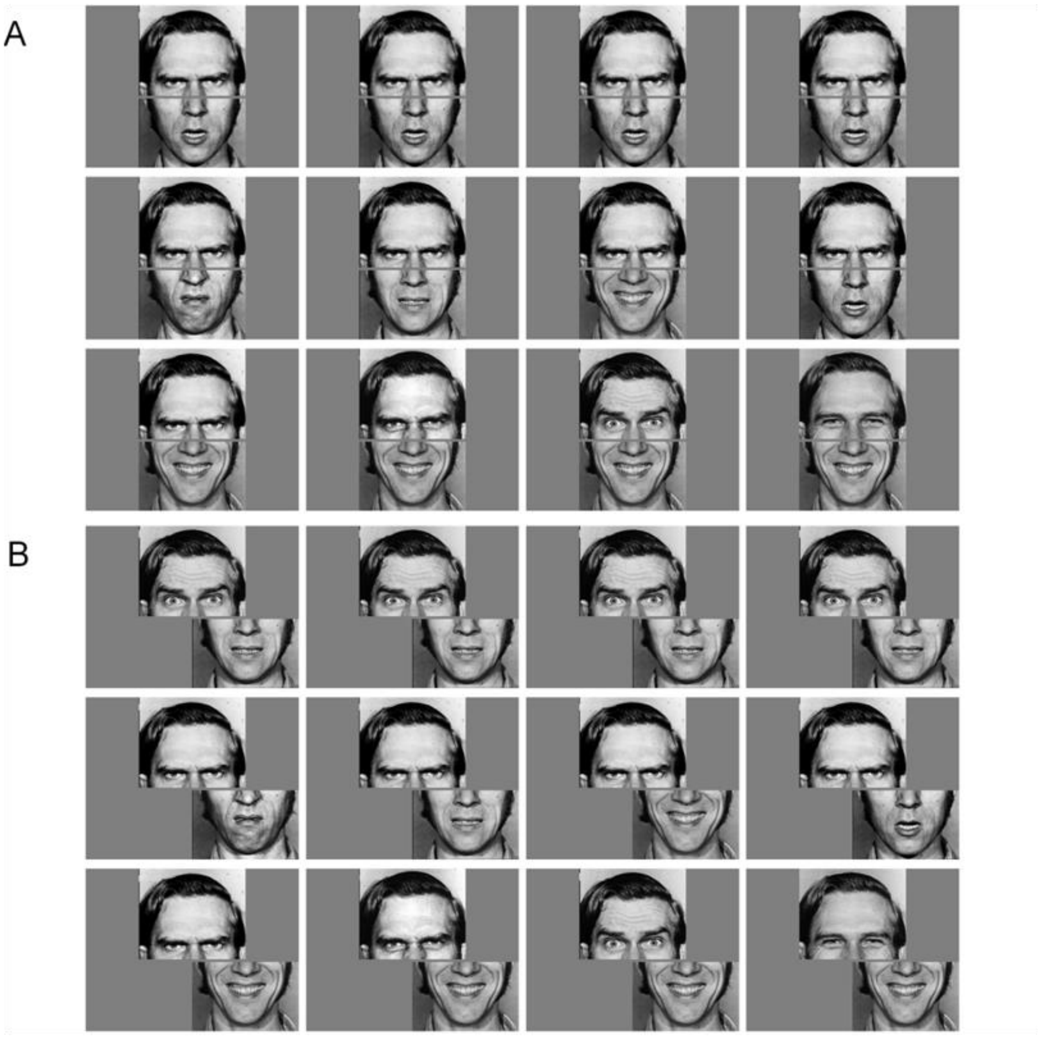


Figure 3.1. Examples of experimental stimuli used to create trial blocks in Experiment 2. A) Aligned conditions (top row: no change, middle row: bottom change, bottom row: top change); B) Misaligned conditions (top row: no change, middle row: bottom change, bottom row: top change). The stimuli used in Experiment 1 involved sequentially presented pairs of images from each of the 6 types of trial block. Note that a small gap between the top and bottom halves of each stimulus emphasises where the parts are joined, even for the aligned images (cf. Rossion, 2013).

There were equal numbers of aligned and misaligned blocks, and the positioning of the image parts in the misaligned blocks was counterbalanced so that half were misaligned to the left, and half to the right. There were 48 blocks in total (6 conditions, repeated 8 times). For the 8 repetitions of each condition, there were 4 blocks for each of the identities used. Within these 4 blocks, each expression was used once as the top half. This meant that within each condition, each identity and expression combination was presented once.

All images were back-projected onto a screen inside the bore of the scanner, approximately 57cm from the participants' eyes. Images were presented in 6 second blocks; this overall block duration is equivalent to those used in other recent studies from my lab, of neural responses to facial expression (Mattavelli et al., 2014; Psalta et al., 2014). Each block contained 4 images, with each image being presented for 750ms with a 750 ms grey screen ISI. There was a 9 second grey screen between each of the blocks. Each stimulus condition was repeated 8 times to give a total of 48 blocks. Hence each scan lasted 12 minutes in total. Images within a block were all derived from the same identity, and the use of each of the 2 identities (models) was randomised across the experiment. Participants monitored all images for the presence of a small red dot (6 pixels in width) that was superimposed at the fixation point on 1 image in each block. Participants were required to respond, with a button press, as soon as they saw the image containing the target red dot. Images for the experimental scan were presented using PsychoPy2 (Peirce, 2007).

3.3.3.2 Imaging Parameters

All scans were conducted using a GE Signa HDx 3T MRI system (General Electric, Waukesha, WI, USA) with an eight channel phased array head coil (MRI Devices Corp., Gainesville, FL). Data were acquired using a gradient echo planar imaging (EPI) sequence with acquisition parameters: 38 contiguous axial slices, repetition time (TR) 3 seconds, echo time (TE) 32.5 milliseconds, flip angle 90°. The field of view (FOV) was 28.8 x 28.8 cm with an acquisition matrix of 128 x 128 and slice-thickness of 3mm, giving a voxel size of 2.25 x 2.25 x 3mm. A T1-weighted Fluid-Attenuated Inversion Recovery (T1-FLAIR) volume was acquired with the same slice orientation and slice thickness with an acquisition matrix of 512x512, giving an in-plane resolution of 0.5625x0.5625mm. To improve registration, the EPI image was initially co-registered with the high resolution initial structural image (T1-weighted FLAIR) containing the same number of slices as the EPI scan before being registered to the high resolution main structural scan (T1-weighted, 1.13 x 1.13 x 1 mm) for each participant. This was then co-registered to the standard MNI 152 brain.

3.3.3.3 fMRI Analysis

Analysis was conducted using FEAT v 5.98 (http://www.fmrib.ox.ac.uk/fsl). The initial 9 seconds of each scan were removed from the analysis to allow T1-saturation effects to subside. Motion correction (McFLIRT; FSL) was applied followed by spatial smoothing (Gaussian, Full Width at Half Maximum 6 mm) and temporal high-pass filtering with a cut off of 0.01 Hz. Face-selective regions were defined in each individual from the functional localiser by using the average of the face > place and face > scrambled face contrasts. The combined statistical maps were thresholded at $p < .01$ (uncorrected). For each individual, the OFA, FFA and pSTS were identified by contiguous clusters of voxels activated above threshold from the above contrast in posterior occipital cortex, inferior fusiform gyrus and superior temporal lobe.

For each individual, the time series of the filtered MR data for each voxel from the experimental scan within each functionally localised ROI was converted to percentage signal change. These were then averaged to produce the time series for each participant within each ROI for each of the experimental conditions. The individual time series data were normalised by subtracting each time point by the zero point at the beginning of the block. These data were then averaged across participants to give the overall mean time series for each condition. The peak response to each condition was taken as the average of TR 2 and TR 3 (corresponding to 6 and 9 seconds after stimulus onset). These peak responses were then entered into repeated measures ANOVAs to determine significant differences between conditions for each ROI.

The primary focus of interest was in neural responses from pSTS based on a functional localiser applied at the individual participant level. However, to determine whether other regions might demonstrate a holistic response to expressions, a whole brain analysis was performed in which the behavioural data from Experiment 1 were used as regressors. A box car function was defined modelling all blocks in the scan run, with each block weighted by the mean RT of that condition. This was convolved with a single gamma hemodynamic response function and then regressed against the BOLD response at each voxel. The resulting statistical maps for each individual were combined using a higher-level mixed effects analysis (FLAME, FSL). The combined statistical maps were thresholded at $z > 2.8$, $p < .05$ (cluster corrected). This process was then repeated using the % error data as a regressor.

3.4 Results

3.4.1 Experiment 1

The aim of Experiment 1 was to demonstrate the facial expression composite effect with the stimuli and presentation times to be used in the fMR-adaptation study. There were 6 conditions involving aligned or misaligned pairs with no change between the images, a bottom half change, or a top half change. Participants monitored the top half of pairs of face images to detect whether the facial expression in the top half remained the same, or was different across the two faces.

First the accuracy of responses were measured when judging whether the top half of each image was the same or different. As participants were asked to make their judgements based only the top half of each image, the correct responses in each condition were 'same' for no change pairs, 'same' for the bottom change pairs, and 'different' for the top change pairs. Percent correct responses were calculated for each condition for each participant, and then averaged across all participants to give an overall percent correct response measure. The data are displayed as percentage errors in Figure 3.2A to facilitate comparison with reaction times shown in Figure 3.2B.

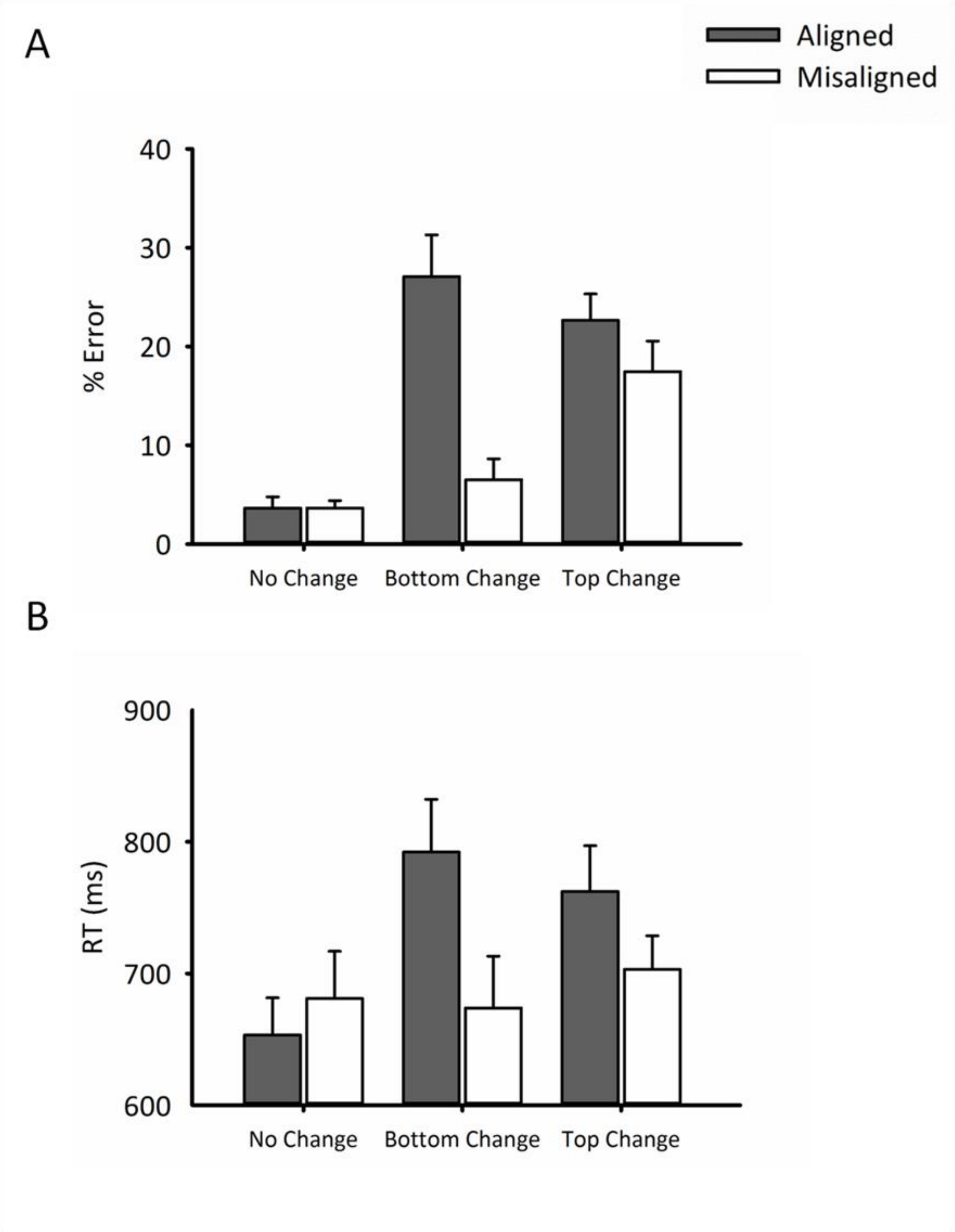


Figure 3.2. (A) Percent error responses for the same-different task in Experiment 1. The critical result is reduced performance (increased errors) in the bottom change condition compared to the no change condition when the stimuli are aligned, but not when they are misaligned, even though in all 4 of these conditions the top halves of the stimuli are to be judged 'same' – the facial expression composite effect. The top change condition is less important because it involves a change in correct response (now 'different' instead of 'same'). (B) Median response times for the same-different task in Experiment 1. RTs were longer for the bottom change condition compared to no change condition when the stimuli were aligned, but not misaligned – again demonstrating the expression composite effect.

The proportion of correct responses was entered into a 2 x 3 repeated measures ANOVA with the factors Alignment (aligned, misaligned) and Condition (no change, top change, bottom change). The ANOVA showed a significant effect of Alignment (F(1,15) = 38.37, p < .001) and Condition (F(2,30) = 19.48, p < .001). Bonferroni pairwise comparisons demonstrated that the effect of Alignment was driven by more accurate responses in the misaligned versus aligned conditions (p < .001). The effect of Condition was driven by more accurate responses in the no change versus top change (p < .001) and bottom change (p = .001) conditions. However, these main effects were both qualified by the presence of a significant Alignment x Condition interaction (F(2,30) = 10.82, p < .001). Paired t-tests demonstrated this was a result of lower accuracy in the bottom change condition when the stimuli were aligned, compared to misaligned (t(15) = -5.54, p < .001) but no difference between the no change aligned and misaligned conditions (t(15) = .432, p = .672). This part of the interaction is the critical test of the facial composite effect, because in all four of these conditions participants were making equivalent responses (that the top halves were the 'same'). In addition, there was also a non-significant trend demonstrating lower accuracy for the top change condition when the stimuli were aligned, compared to misaligned (t(15) = -1.86, p = .083). This may reflect a slight interference from the bottom half of the face when the two halves are aligned, giving slightly higher errors.

Response times to each condition were also measured. Median RTs were taken for each condition, for each participant and an overall median RT was calculated for each condition across all participants (Figure 3.2B). These median RTs were entered into a 2 x 3 repeated measures ANOVA with the factors Alignment (aligned, misaligned) and Condition (no change, top change, bottom change). This ANOVA demonstrated significant main effects of Alignment (F(1,15) = 18.24, p = .001) and Condition (F(2,30) = 16.36, p < .001). Bonferroni pairwise comparisons demonstrated the effect of Alignment was driven by longer RTs when the stimuli were aligned, compared to misaligned (p = .001) and the effect of Condition was driven by a longer RT in both top change (p = < .001) and bottom change (p < .001) conditions relative to no change.

Again, interpretation of these main effects needs to be qualified by a significant Alignment x Condition interaction (F(2,30) = 11.62, p < .001). Paired t-tests demonstrated this was due to longer response times in the aligned versions of both top change and bottom change conditions when compared to their misaligned counterparts (bottom change: t(15) = 4.69, p < .001, top change: t(15) = 3.04, p < .001). No difference was seen in the response times between the aligned and misaligned versions of the no change condition (t(15) = -1.54, p = .145). Paralleling the analysis of accuracy data, the slower response times in the aligned compared to misaligned version of the bottom change condition, and the lack of difference in response time for the no change condition, illustrate the key components of the face composite effect.

In sum, behavioural results from the RT and accuracy data show the facial expression composite effect where participants find it more difficult to judge the top half of the images as the same when the bottom half is changing and the two halves of each image are aligned into an overall facial configuration, compared to when they are in a misaligned form.

3.4.2 Experiment 2

The aim of this experiment was to investigate properties of the right pSTS response to facial expressions, using conditions comparable to those in the behavioural Experiment 1. The principal focus of the analysis was pSTS because of its hypothesised role in facial expression perception in the leading neural model of face perception, (Haxby et al., 2000), and on right rather than left pSTS because right pSTS is more reliably identified at the individual participant level with the functional localiser scan and has therefore been targeted in previous studies (Harris et al., 2012, 2014). To parallel Experiment 1, there were 6 different types of block in the experimental scan, involving aligned or misaligned pairs with no change between the images, a bottom half change, or a top half change

In order to check whether participants were watching the top halves of the stimuli throughout the experiment, as instructed, they were given the task of pressing a response button every time they saw a small red dot presented at the fixation point. Performance on this red dot detection task was high, with a mean accuracy of 99% correct responses and mean RT of 447ms. To confirm that there were no differences in overall attentional demands between aligned and misaligned stimuli, the average response times to aligned and misaligned conditions for each participant were entered into a paired t-test. There was no significant difference in response times to the red dot, t(21)=1.39, p = .18.

The pSTS, FFA and OFA were localised in the left and right hemispheres using the independent functional localiser scan. The OFA and FFA could be identified in both the left and right hemispheres for 23/26 participants. In contrast to the OFA and FFA, the pSTS could be reliably identified in the right hemisphere of 22/26 participants, but in the left hemisphere for only 15/26 participants. Therefore, only the region localised in the right hemisphere was used for the pSTS. This relatively poor face responsiveness of left pSTS may be due to its possible role in more audiovisual integration of vocal and facial speech signals (Calvert, 2001; Pelphrey et al., 2005; Wright et al., 2003). Average MNI coordinates and number of voxels for each localised ROI are provided in Table 3.1.

Table 3.1. Average MNI coordinates in mm (mean and SE), size in voxels, and number of participants where the region could be identified, for each ROI.

ROI	Coordinate			No. of Voxels	No. of Participants
	x	y	z		
Right OFA	41 ± 1	-80 ± 2	-15 ± 1	187	26
Left OFA	-41 ± 1	-83 ± 1	-14 ± 1	107	23
Right FFA	41 ± 1	-56 ± 1	-23 ± 1	223	26
Left FFA	-40 ± 1	-60 ± 2	-23 ± 1	114	23
Right pSTS	51 ± 1	-61 ± 2	1 ± 1	110	23

There was no effect of hemisphere for the OFA (F(1,22) = 0.16, p = .696) or FFA (F(1,22) = 1.58, p = .221), so the data from the left and right hemispheres of these regions were combined. In terms of Haxby et al.'s (2000) neural model of face perception, results for the pSTS and FFA are the most instructive, as these lie on separate neural pathways considered to be critically involved in the perception of expression (pSTS) or to be involved in other aspects of face perception (FFA). The OFA was considered as of less interest because it lies on both neural pathways in Haxby et al.'s (2000) model, but data from the OFA were analysed, for completeness.

First, the time series data for each participant was averaged across participants to give an overall mean time series for each condition, for each ROI (Figure 3.3). The peak responses in the right pSTS, which form this chapter's principal focus of interest (Figure 3.3, panel A) were then analysed. A 2x3 ANOVA with the factors Alignment (aligned, misaligned) and Condition (no change, bottom change, top change) demonstrated a significant effect of Condition (F(2,44) = 7.62, p = .001), but not of Alignment (F(1,22) < 1). The Alignment x Condition interaction was not significant (F(2,44) < 1). The effect of Condition was driven by a smaller peak percentage signal change in the no change condition compared to both the bottom change (t(22) = -3.75, p = .001) and top change conditions (t(22) = -2.93, p = .008), with no difference between the signal change in the bottom and top change conditions (t(22) = .301, p = .797). This pattern is consistent with a feature-based response, with no evidence of the critical interaction between Alignment and Condition that would demonstrate holistic perception.

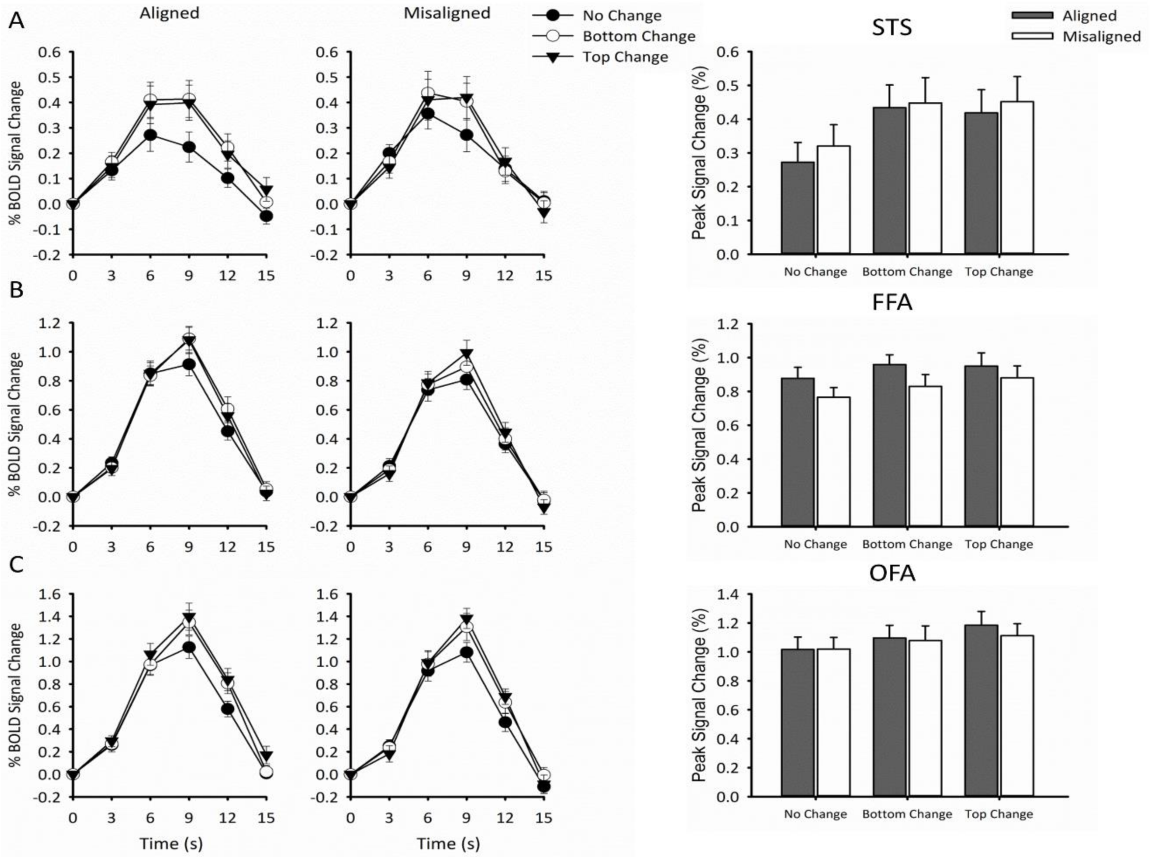


Figure 3.3. Overall mean MR time series for each condition for aligned and misaligned stimuli, and peak % BOLD signal change for right pSTS (row A), FFA (row B) and OFA (row C). Analysis of the responses in right pSTS revealed a smaller peak response in the no change condition compared to both the bottom change (p = .001) and top change conditions (p = .008), with no difference between the bottom and top change conditions. This pattern held for aligned and misaligned stimuli. In FFA, there was only a main effect of Alignment, with a higher peak response to aligned than misaligned stimuli (p = .021). Error bars represent standard error of the mean.

The FFA showed a different pattern of results to the pSTS (Figure 3.3, panel B). A 2x3 ANOVA showed a significant effect of Alignment (F(1,25)= 6.11, p = .021), but only a borderline effect of Condition (F(2,50)= 2.56, p = .088). The Alignment x Condition interaction was not significant (F(2,50) < 1. The effect of Alignment was driven by a significantly higher peak percent signal change to the aligned compared to misaligned stimuli (t(25) = 2.47, p = .021).

The OFA did not produce any findings that reached conventional levels of statistical significance (Figure 3.3). There was no effect of Alignment (F(1,25) < 1, and after Greenhouse-Geisser correction for a violation of sphericity ($\chi^2(2) = 9.03$, p = .011) only a borderline effect of Condition (F(1.523,38.07) = 3.32, p = .059). There was no Alignment x Condition interaction (F(2,50) < 1).

To determine if other regions showed a holistic response, a whole brain analysis was also conducted. The % error and response time data from Experiment 1 were used as regressors to identify regions that might show a holistic response. The resulting group statistical parametric map identified 2 clusters of activity, in the right inferior frontal gyrus (IFG) and in the right fusiform gyrus. Table 3.2 shows the peak voxel intensity, co-ordinates and size of the ROIs based on the % error and RT data.

Table 3.2. Peak intensity and MNI coordinates (mm) for maximally active voxel, and size in voxels for each ROI identified using the mean RT and % error data from Experiment 1 as a regressor.

ROI	Peak Intensity (z score)	Coordinate			No. of Voxels
		x	y	z	
% Error					
Right Fusiform	4.86	38	-50	-22	771
Right IFG	3.90	48	4	18	411
RT					
Right Fusiform	4.97	40	-50	-24	656
Right IFG	4.09	48	6	18	654

These data were used to create masks of the regions identified (right fusiform, and right IFG). The time series data for each participant was then taken and averaged across participants to give an overall mean time series for each condition, for each ROI. The peak responses for each condition for each ROI were then calculated. As can be seen from Table 3.2, the peak intensities were very similar for both the ROIs identified using the RT and % error data. This was also reflected in the peak response to each individual condition, therefore only the % error regressor data is presented for illustration purposes, in Figure 3.4. The right IFG shows the classic pattern demonstrated in the expression composite effect – a higher response to bottom change when the face is aligned, compared to when misaligned. It also shows a smaller response to the no change compared to the change conditions. In contrast, the fusiform gyrus shows a more general overall difference in responsiveness between aligned and misaligned images. This is consistent with the known involvement of fusiform cortex in the holistic perception of faces (Kanwisher et al., 1997; Andrews et al., 2010), but does not imply holistic processing of expression per se.

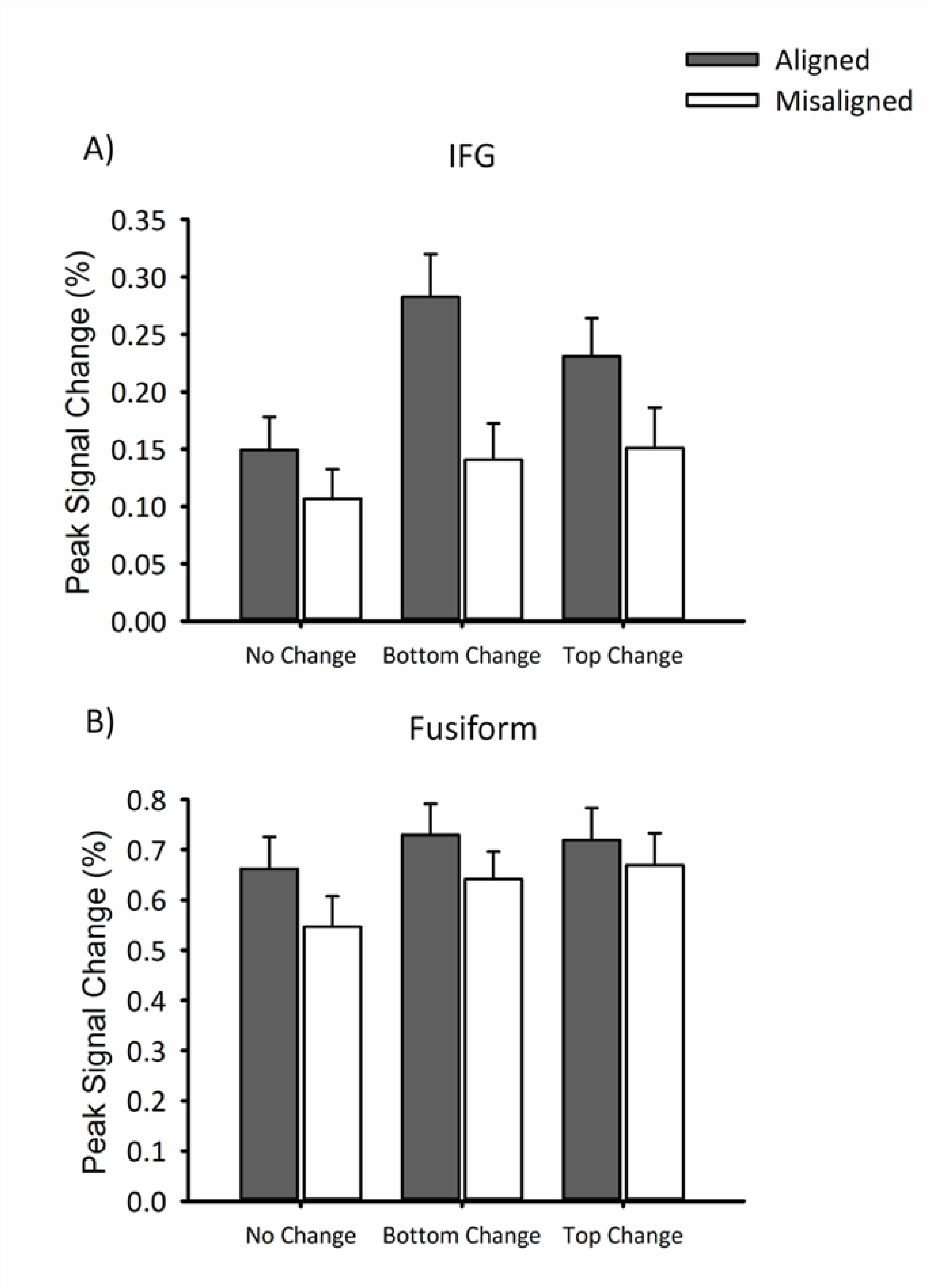


Figure 3.4. Overall mean peak % BOLD signal change for each condition for aligned and misaligned stimuli for the right IFG (A), and right fusiform (B). Regions defined using the % error data from Experiment 1 as a regressor. Error bars represent standard error of the mean.

3.5 Discussion

In this chapter, an fMR-adaptation paradigm was used to investigate neural responses to facial expressions in core regions of Haxby et al.'s (2000) neural model of face perception, focussing particularly on pSTS because of its hypothesised role in the perception of expression. By using a no change condition as a baseline promoting maximal adaptation, it was possible to demonstrate release from adaptation in right pSTS to conditions in which changes in expression were located in the upper or lower parts of the stimuli. This shows that the right pSTS was encoding such changes, even though the incidental task of detecting a red spot was irrelevant to perceiving the facial expression. Moreover, the degree of adaptation in right pSTS was equivalent whether the changes occurred in the fixated, task-relevant (top half) or non-fixated (bottom half) part of each stimulus.

This pattern of neural response in pSTS was the same regardless of whether the top and bottom parts of the stimuli were aligned into a face-like overall configuration, or misaligned by offsetting the parts to make the overall image less face-like. The contrast between aligned and misaligned variants of the stimuli is of theoretical importance, as it is now widely used to probe holistic processing of faces in studies of the perception of face identity and facial expression (Calder et al., 2000; Maurer, Le Grand, & Mondloch, 2002; Rossion, 2013; Young, Hellawell, & Hay, 1987). The logic underlying the contrast is that holistic processing of the stimulus as a face is only possible when the constituent parts are correctly aligned, and that a consequence of holistic processing will be to enhance perceived differences between stimuli that share common parts - for example, making the top change stimuli look more different from each other when in the aligned than in the misaligned arrangement. This enhanced perception of differences between aligned than misaligned stimuli was demonstrated behaviourally in Experiment 1, so it is noteworthy that the results do not show such an effect in the neural responses from pSTS. Instead, it seems that pSTS is sensitive to any change in face parts (with a release from adaptation in both top change and bottom change conditions) but does not require that the stimulus is particularly face-like (as shown by the equivalent release from adaptation across aligned and misaligned stimuli). This complements Harris et al.'s (2012) finding that pSTS responds more or less linearly to all changes in facial features that communicate emotion.

A possibility that needs to be considered is that the differences in the pattern of the results between the behavioural (Experiment 1) and fMRI (Experiment 2) data might reflect task differences. In the behavioural experiment, participants were asked to detect changes in facial expression. In contrast, in the fMRI experiment, participants were asked to detect a red dot superimposed on some of the faces. An explicit holistic task was not used in the fMRI experiment because the aim was to examine how facial expression is encoded irrespective of task difficulty. Using an explicit task of holistic processing would introduce differences in task difficulty across conditions and as a result, produce attentional differences across conditions. Therefore it was important to use a task independent of the experimental manipulation to ensure all stimuli were attended to equally in the fMR experiment. Since the expression composite effect is considered to reflect mandatory holistic face perception and no previous work has suggested that it is affected by the task, this offered the best way to eliminate potential attentional confounds. It is also important to note that the facial *identity* composite effect can be demonstrated using a similar fMRI experimental procedure (Schiltz and Rossion, 2006) to that presented here.

The FFA showed a different pattern of response than pSTS, with the only finding that reached the conventional level of statistical significance being a main effect of alignment, with higher overall response to aligned than to misaligned stimuli. These results are consistent with previous studies that used fMR-adaptation with composite faces to reveal a holistic response to facial identity in the FFA (Andrews, Davies-Thompson, Kingstone, & Young, 2010; Schiltz, Dricot, Goebel, & Rossion, 2010; Schiltz & Rossion, 2006). The pattern is also consistent with Kanwisher et al.'s (1997) landmark study defining the properties of the FFA, which found a stronger response to normal faces than to scrambled arrangements of face parts, as misaligning the stimuli can be considered a simple variant of face scrambling. This finding reveals that there are fundamentally different neural representations of faces in the FFA and pSTS. The representation in the FFA is sensitive to the correct configuration of the facial features, whereas the pSTS appears to encode facial features independently.

To determine if regions outside the core face-selective regions showed a holistic response to facial expression, a group analysis was performed. This analysis used the behavioural data from Experiment 1 as a regressor, as this had shown a holistic response to expression. The independence of the behavioural (Experiment 1) and fMRI (Experiment 2) data used in this analysis offers a strong test of whether a region can be linked to a specific pattern of responses. This group analysis identified the right fusiform gyrus and right inferior frontal gyrus (IFG) as regions that covaried with behavioural responses. Inspection of the data shown in Figure 3.4 suggests that the fusiform activity was due to a more general holistic response to faces per se, in the form of a higher overall response to all aligned than misaligned stimuli, as had also been shown from the analysis of the FFA defined with the individually-based functional localiser. In contrast, the IFG showed a pattern of response which was more consistent with a holistic response to facial expression, as evidenced by the similarity between the pattern of BOLD responses in IFG (Figure 3.4) and the RTs and errors in the behavioural task (Figure 3.2). These results are consistent with previous studies which have shown that right IFG is part of the extended face processing network (Ishai et al. 2008; Davies-Thompson et al., 2012) and is involved in the processing of facial expressions (Ishai, Schmidt & Boesiger, 2005; Carr et al. 2003; Dapretto et al., 2006).

It is important to note that in this chapter, the response across all facial expressions was looked at. Although the design does not allow for the data to be explored in this way, it would be interesting to look at the response for each individual expression. This would be particularly interesting as some facial expressions are more recognisable from their bottom halves, and some from their top halves (Calder et al. 2000).

To summarise, it has been shown that right pSTS is sensitive to changes in the facial features that convey emotion regardless of whether these changes occur in the fixated parts of the image or not, and regardless of whether image parts are arranged in a more or a less face-like configuration. Therefore, based on these results, the pSTS cannot be considered the neural locus of the facial expression composite effect. Nonetheless, these findings are consistent with Haxby et al.'s (2000) view that pSTS is an important region in the perceptual analysis of facial expressions and uncover something of this region's *modus operandi*, showing in particular that it is very responsive to changes in expressive features whether or not these form a face-like overall configuration.

This chapter has shown the role of the pSTS in the perception of changeable aspects of the face, such as facial expression. However, there are different types of changes in the face and they can convey different meanings to the observer. This suggests that they might be represented differently in areas such as the pSTS. This will be addressed in the next chapter.

Chapter 4 – Patterns of Response to Changes in Expression and Changes in Viewpoint in Face-Responsive Regions of the Human Brain

4.1 Abstract

Changeable aspects of the face play an important role in social communication. Different types of changeable aspects of faces can convey distinct social information. For example, changes in facial expression indicate how a person is feeling, whereas changes in viewpoint indicate their direction of attention. Models of face processing propose that the posterior superior temporal sulcus (pSTS) plays a central role in the processing of changeable aspects of faces (Haxby et al., 2000). However, it is not clear whether different types of change have distinct neural representations. This chapter addressed how two distinct types of facial change (expression and viewpoint) are represented in face-selective regions. fMRI was used to investigate the neural representation of facial expression and viewpoint. In Experiment 1, participants viewed sequences of faces that varied in either expression or viewpoint. Using MVPA, distinct patterns of response for changes in expression and changes in viewpoint were found within the face-responsive regions of the OFA, FFA, STS. Distinct patterns of response to expression were found in the IFG and amygdala. The patterns of response to expression and viewpoint were largely invariant to changes in facial identity and these patterns of response were consistent across participants. In Experiment 2, an fMR-adaptation paradigm was used to examine the selectivity of the ROIs to changes in expression and viewpoint. Participants viewed images of faces that changed in expression and viewpoint. The OFA, FFA, pSTS demonstrated release from adaptation to both expression and viewpoint whereas the IFG and amygdala only showed selectivity for expression. Together these results show distinct patterns of response to expression and viewpoint changes in face-responsive regions of the human brain that are consistent with the distinct social information conveyed by these different types of change.

4.2 Introduction

Chapter 3 explored the representation of facial expression within face-selective regions. The aim of this chapter was to ask how different types of changeable aspects of faces are represented in these regions. Specifically this chapter addresses whether there are distinct representations for facial expression and viewpoint.

Changeable aspects of the face convey a wealth of socially important information (Bruce and Young, 2012). These changes are often considered to be of two main types: non-rigid and rigid. Non-rigid changes of the facial muscles give rise to different facial expressions. In contrast, rigid changes of the head create upward or downward movements (often interpreted as indicating dominance or submissiveness) and rotational movements (that usually signal a shift in a person's focus of attention). Although these changes are often subsumed under the broad heading of changeable aspects of faces, their implications for the perceiver are quite different.

Changeable aspects of the face are thought to be processed independently of invariant facial properties such as identity (Andrews & Ewbank, 2004; Bruce & Young, 2012; Gobbini & Haxby, 2007; Haxby, Hoffman, & Gobbini, 2000). A variety of evidence has shown that the superior temporal sulcus (STS) is important for the processing of facial changes (e.g. Allison et al., 2000). The response of the STS has been shown to be selective for head direction or eye position (Perrett et al., 1985; Pelphrey et al., 2004; Calder et al., 2007; Fang et al., 2007; Natu et al., 2010; Lee et al., 2010; Carlin et al., 2011, 2012), but also for differences in facial expression (Baseler, Harris, Young, & Andrews, 2013; Engell & Haxby, 2007; Harris, Young, & Andrews; Winston, Henson, Fine-Goulden, & Dolan, 2004). Information about these aspects of faces is thought to be relayed from the pSTS to an extended face processing network (Haxby et al., 2000) that includes the amygdala (Morris et al., 1996; Winston et al., 2003) and parts of the frontal and parietal lobes (Calder & Nummenmaa, 2007; Wicker, Michel, Henaff, & Decety, 1998).

Although neural responses to changeable aspects of faces, such as expression and viewpoint, occur in these face-selective regions, it is not clear whether there are distinct representations for these different changes in the face. Changes in facial expression and changes in viewpoint (head direction), convey quite different social signals. They also produce very different changes in the image. Facial expressions result in changes that primarily affect the shape of the internal features of the face (Bruce & Young, 2012), whereas changes in viewpoint produce changes in the pattern of luminance across large regions of the image by exposing and occluding different parts of the face and head. Despite these large visual differences, changeable aspects of faces are widely considered to be processed by the same pathway (Bruce and Young, 1986, Haxby et al., 2000).

The aims of this chapter were to investigate whether there are distinct neural representations for the processing of changes in facial expression and viewpoint, and to evaluate whether these representations are consistent across different individual participants. To achieve these aims, an analysis was performed in MNI space to identify face-responsive regions at the group level that corresponded to the core face processing regions identified by Haxby et al. (2000). In Experiment 1, multi-voxel pattern analysis (MVPA) of fMRI data was used to quantify similarities across participants in the patterns of neural response in the face-responsive ROIs elicited by changes in facial expression or changes in viewpoint. In addition, a regression analysis was used to determine whether these patterns of response were invariant to changes in facial identity. In Experiment 2, an fMR-adaptation paradigm was used to determine the selectivity of response to expression and viewpoint across the face-responsive regions.

Given the different types of information conveyed by the two changeable aspects of faces addressed by this chapter, it is predicted that these will have a distinct representation within face-selective regions. Specifically, it is predicted that there will be distinct patterns of response to changes in expression, and distinct patterns of response to changes in viewpoint. In particular, this should be evident in the pSTS which is the face-selective region implicated in the processing of changeable aspects of faces. However in the extended regions of the face-processing system this distinction may be less clear. In regions where higher-order processing is occurring, for example the IFG, specific combinations of expressions and viewpoints might have more relevance in these regions and therefore a distinct representation is less likely.

4.3 Method

4.3.1 Participants

In Experiment 1, data were analysed from the functional localiser scans acquired as part of previous fMRI studies (Harris, Young, & Andrews, 2012; Mattavelli et al., 2014; Psalta, Young, Thompson, & Andrews, 2014). There were 83 right-handed participants (49 females, mean age 24.1 ± 5.8). There were 31 participants in Experiment 2 (17 female, mean age 23.5 ± 3.3). All participants had normal or corrected-to-normal vision with no history of neurological conditions. Written consent was obtained from all participants and the studies were approved by the York Neuroimaging Centre Research Ethics Committee.

4.3.2 Stimuli and Design

4.3.2.1 Experiment 1

Changes in expression and viewpoint that result from movements of the face are clearly conveyed from the apex of the movement (Bruce & Young, 2012). Therefore, in line with the majority of previous studies of gaze and expression, static images were used to achieve a high level of experimental control across conditions. A block design was used with five stimulus conditions. There were four face conditions (Figure 4.1): (1) *same-identity, different-expression* (sIdE), (2) *different-identity, different-expression* (dIdE), (3) *same-identity, different-viewpoint* (sIdV), (4) *different-identity different-viewpoint* (dIdV). Face images were taken from the Radboud Faces Database (Langner et al., 2010). *Same-identity* conditions had the same facial identity within each block. *Different-identity* conditions had different identities in each block. *Different-expression* blocks showed frontal images posing five different expressions. *Different-viewpoint* conditions showed five sequential images of left-profile, left ¾, frontal, right ¾, right profile faces. The final stimulus condition (used to localise face-responsive regions) consisted of images taken from each face category that were phase-scrambled in the Fourier domain. The five expressions used were fear, anger, happy, sad and disgust. These are the five expressions considered to be basic emotions, described in Chapter 1, section 1.2.3. *Facial expressions*.

Figure 4.1. Examples of stimuli from the different conditions. Each row shows the sequence of images in a representative stimulus block from the different conditions.

Images (approximately 6°x8°) were back-projected onto a screen inside the bore of the scanner, approximately 57 cm from the participants' eyes. Images were presented in 6 second blocks. Each block contained 5 images, with each image being presented for 1000 ms with a 200 ms black screen inter stimulus interval (ISI). These image parameters were used to make sure that there was no perception of apparent motion during the sequence. There was a 9 second grey screen between each of the blocks. Each stimulus condition was repeated 5 times to give a total of 35 blocks. Each scan therefore lasted 9 minutes in total. To ensure attention throughout the scan, participants were instructed to monitor each image for the presence of a red dot that was superimposed randomly on 1 image per block. Participants were required to respond, with a button press, as soon as they saw the image containing the red dot target. The target could appear in any location on the image and was counterbalanced (in terms of trial and location) across conditions. Accuracy on the red dot task was high (98.6 ± 5.0%), indicating that participants paid attention to the sequence of images.

4.3.2.2 Experiment 2

This experiment used an fMR-adaptation design, with four stimulus conditions (Figure 4.2): 1) same-expression, same-viewpoint (sEsV), 2) different-expression, same-viewpoint (dEsV), 3) same-expression, different-viewpoint (sEdV), 4) different-expression, different-viewpoint (dEdV). Same-expression conditions had the same facial expression within each block, but different expressions (fear, anger, disgust, sad, happy) were used in different blocks. Different-expression conditions had all five different expressions in each block (fear, anger, disgust, sad, happy). Same-viewpoint conditions had the same viewpoint in each block, but different viewpoints (left-profile, left ¾, frontal, right ¾, right profile) were used in different blocks. Different-viewpoint conditions had all five different viewpoints in each block (left-profile, left ¾, frontal, right ¾, right profile). As the patterns of response to changes in expression and changes in viewpoint in Experiment 1, were found to be invariant to changes in identity (see results, section 4.4.1. *Experiment 1*), the identify factor was removed from Experiment 2 to allow a more specific focus on expression and viewpoint. Therefore, faces in each block all had the same identity, so they only varied in viewpoint or expression.

Figure 4.2. Examples of stimuli from the experimental conditions for experiment 2. Each row shows the sequence of images in a representative stimulus block from the different conditions.

Images (approximately 6°x8°) were back-projected onto a screen inside the bore of the scanner, approximately 57 cm from the participants' eyes. Images from each condition were presented in 6 second blocks, with each block containing 5 images. Each image was presented for 1000 ms with a 200ms grey screen ISI. There was a 9 second inter-block interval during which a grey screen with a white fixation cross was presented. Each stimulus condition was repeated 10 times giving a total of 40 blocks. The participants' task during the scan was to monitor images for the presence of a red dot. Accuracy on this task was high, with a mean accuracy of 99.0% (SD 2.2) demonstrating the participants were maintaining attention to the stimuli throughout.

4.3.3 Imaging Parameters

All scans were conducted using a 3 Tesla MRI system with an eight channel phased array head coil (GE Signa Excite HDx 3.0T, High resolution brain array, MRI Devices Corp., Gainesville, FL) tuned to 127.4MHz. A gradient echo planar imaging (EPI) sequence was used to acquire the data. The acquisition parameters were: 38 contiguous axial slices, repetition time (TR) 3 seconds, echo time (TE) 32.5 milliseconds, flip angle 90°, field of view (FOV) 28.8 x 28.8 cm, matrix 128 x 128, slice-thickness 3mm, voxel size 2.25 x 2.25 x 3mm. To improve registration, the EPI image was co-registered with a T1-weighted image taken in the same plane, before being registered to the high resolution main structural scan (T1-weighted, 1.13 x 1.13 x 1 mm) of each participant. Linear affine transformations were calculated to align session data to intermediate, high resolution and standard (MNI) anatomical spaces using FLIRT (Jenkinson et al., 2001).

4.3.4 fMRI Analysis

4.3.4.1 Experiment 1

For the experimental scan, the initial 9 seconds were removed from the analysis in order to reduce the effects of magnetic stimulation. Motion correction was applied followed by spatial smoothing (Gaussian, full width at half maximum 6 mm) and temporal high-pass filtering (cut off, 0.01 Hz). Individual participant data were then entered into a higher-level group analysis using a mixed effects design (FLAME, http://www.fmrib.ox.ac.uk/fsl) whole-brain analysis. To identifiy face-responsive regions, ROIs were defined by an average of each face condition > scrambled condition. The average statistical map was thresholded at $Z > 3.1$, $p < .001$. The peak voxel was located in areas corresponding to the OFA, FFA, pSTS, IFG and amygdala in each hemisphere. ROIs were identified by contiguous clusters of 500 voxels surrounding the peak voxel, activated above threshold from the above contrast. It was not possible to localise 500 contiguous voxels in the left hemisphere for the IFG and amygdala and so these contained 187 and 333 voxels respectively. In addition to the present chapter, these ROIs were used for the analyses in Chapters 5 and 6.

Pattern analyses were performed as described in Chapter 2, section 1.3.2.3. *LOPO Methods.*

Paired samples t-tests were used to test the difference between the within-condition and between-condition correlations. If a given stimulus category evoked a distinct pattern of activity, then the within-condition correlations (e.g. sIdE – sIdE) for the individual participant and rest of the group data should be higher than the between-condition correlations (e.g. sIdE – sIdV).

To provide a measure of decoding accuracy for changes in viewpoint and expression, data were also run through a k-nearest neighbour classifier (k=1) using correlation as the distance measure. This allowed us to determine how well the patterns for expression and viewpoint could be discriminated. One sample t-tests were then used to test whether decoding accuracy differed significantly from chance across participants.

A regression analysis was used to assess the patterns of response to the four face conditions. For each factor, a binary regressor was generated representing an idealised correlations matrix in which 0 or 1 was entered in different cells. Two regressors were developed, representing patterns of response which would suggest identity invariance, or sensitivity to changes in identity. Each regressor was then entered into a simple linear regression, with the outcomes defined as the correlations matrices obtained from the MVPA concatenated across LOPO iterations. This analysis yielded a beta value for each regressor which would be expected to be significantly greater than zero if that regressor was able to explain a significant amount of the variance in the MVPA data.

4.3.4.2 Experiment 2

The fMRI analysis was conducted using FEAT (http://www.fmrib.ox.ac.uk/fsl). The initial 9 seconds of each scan were removed from the analysis to allow T1-saturation effects to subside. Motion correction (McFLIRT; FSL) was applied followed by spatial smoothing (Gaussian, Full Width at Half Maximum 6 mm) and temporal high-pass filtering with a cut off of 0.01 Hz. Individual participant data were entered into a higher-level group analysis (mixed effects, FLAME http://www.fmrib.ox.ac.uk/fsl). The functional data were first registered to a high resolution T1-anatomical image and this was then co-registered to the standard MNI 152 brain. An fMR-adaptation paradigm was used in order to determine the selectivity of voxels in the face-responsive ROIs, to expression and viewpoint. The following contrasts were used: 1) different expression, different viewpoint > same expression, different viewpoint (dEdV > sEdV), 2) different expression, same viewpoint > same expression, same viewpoint (dEsV > sEsV), 3) different expression, different viewpoint > different expression, same viewpoint (dEdV > dEsV), 4) same expression, different viewpoint > same expression, same viewpoint (sEdV > sEsV). Contrasts 1 and 2 represent a release from adaptation to expression, and contrast 3 and 4 represent a release from adaptation to viewpoint. Statistical maps were thresholded at $Z > 2.3$, cluster corrected, $p < .050$.

4.4 Results

4.4.1 Experiment 1

This chapter asked whether there are distinct neural representations for the processing of changes in facial expression and viewpoint. MVPA was used to investigate the similarity of the patterns of response to the four different stimulus conditions across participants, (1) *same-identity, different-expression* (sIdE), (2) *different-identity, different-expression* (dIdE), (3) *same-identity, different-viewpoint* (sIdV), (4) *different-identity different-viewpoint* (dIdV). Figure 4.3 shows matrices of the correlations between different conditions across the face-responsive ROIs. The results show higher correlations between conditions in which expression changed but not viewpoint, or between conditions in which viewpoint changed but not expression, irrespective of changes in identity. Due to the normalisation process used, the correlation values shown in Figure 4.3. cannot be interpreted literally - the key point to note is the difference between the within-condition and between-condition correlations.

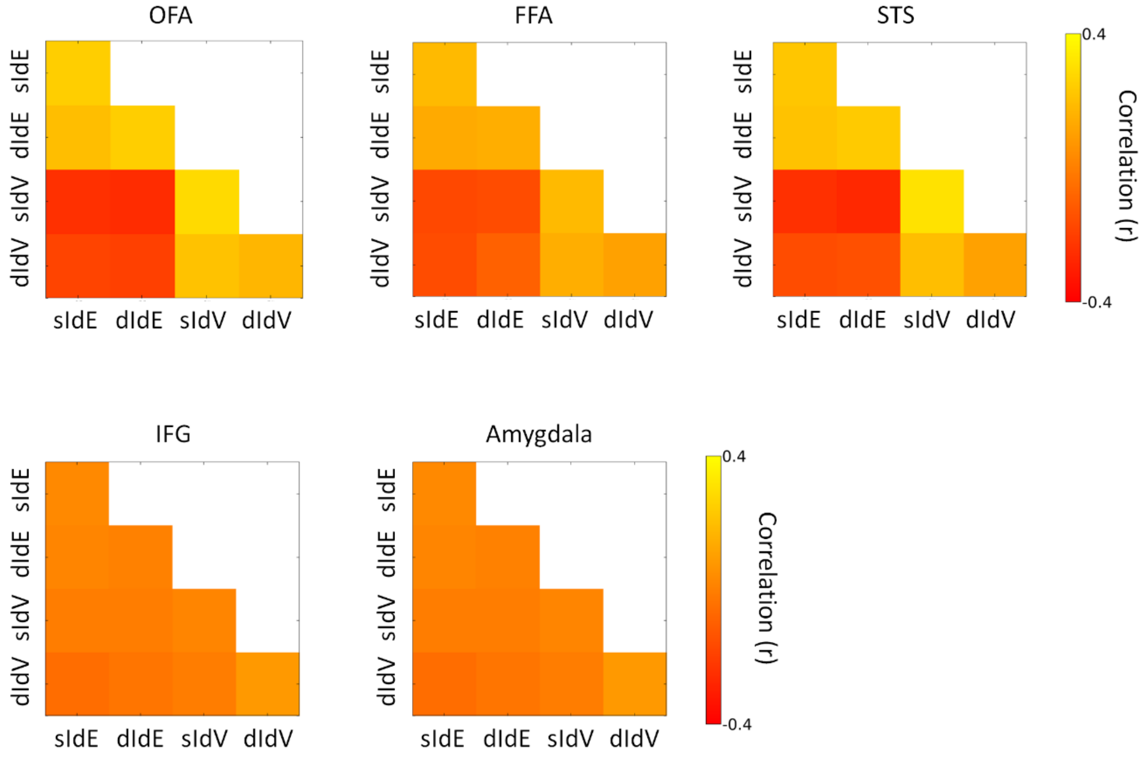


Figure 4.3. Correlation matrices showing similarity in the patterns of response between conditions in each of the ROIs. sIdE: same identity, different expression, dIdE: different identity, different expression, sIdV: same identity, different viewpoint, dIdV: different identity, different viewpoint.

For example, the similarity in the patterns of response to conditions in which both involved a change in expression (sIdE-sIdE, dIdE-dIdE, sIdE-dIdE) was significantly greater than the similarity between patterns of response in which expression changed in one condition and viewpoint changed in the other condition for the core regions (OFA: all t > 6.00, p < .001, FFA: all t > 3.52, p ≤ .001 STS: all t > 3.68, p < .001). The results in the extended regions were less clear, with the IFG showing a high degree of similarity in dIdE-dIdE compared to the similarity between patterns of response in which expression changed in one condition and viewpoint changed in the other condition (t = 3.25, p < .010). The amygdala demonstrated a high degree of similarity between sIdE-dIdE, compared to the similarity between patterns of response in which expression changed in one condition and viewpoint changed in the other condition (t = 4.48, p < .001).

In the same way, the similarity in the patterns of response to conditions which both involved a change in viewpoint (sIdV-sIdV, dIdV-dIdV, sIdV-dIdV) was significantly greater than the similarity between patterns of response in which expression changed in one condition and viewpoint changed in the other condition in the core regions (OFA: all t > 8.13, p<.001, FFA: all t > 4.42, p < .001, STS: all t > 7.67, p < .001). These results also demonstrate the consistency in the patterns of response across individuals, showing that to some extent, different individuals use similar neural coding in order to process expression changes and similar neural coding for viewpoint changes. The patterns of response in the IFG did not show a greater similarity for conditions which both involved a change in viewpoint over conditions which involved a change in expression and viewpoint. However, the amygdala showed greater similarity between sIdV-sIdV and sIdV-dIdV than between conditions where expression changed in one and viewpoint in the other (all t > 2.37, p < .050).

To further analyse the consistency of the patterns of response to viewpoint and expression, the data were collapsed across identity and a k-nearest neighbour pattern classifier was used to provide a measure of decoding accuracy for viewpoint or expression based upon the correlations. Across the core face-selective ROIs (OFA, FFA and STS), the pattern for facial expression could be correctly discriminated on over 84% of trials and a similar level of discrimination was evident for patterns of response to changes in viewpoint. The classification performance for the extended regions was lower, although the classifier was still able to discriminate expression in both regions, however the classifier was only able to discriminate viewpoint in the amygdala (Table 4.1).

Table 4.1. Performance of pattern classification of Expression and Viewpoint for each region (%)

	Expression	Viewpoint
OFA	85***	85***
FFA	73***	74***
STS	84***	83***
IFG	59**	55 *ns*
Amygdala	58*	62***

*** $p < .001$, ** $p < .01$,* $p < .05$

To determine if the patterns of response to changes in expression and changes in viewpoint are sensitive to changes in identity, models were created which represent idealised coding of identity invariance and sensitivity to identity changes (identity sensitive) (Figure 4.4).

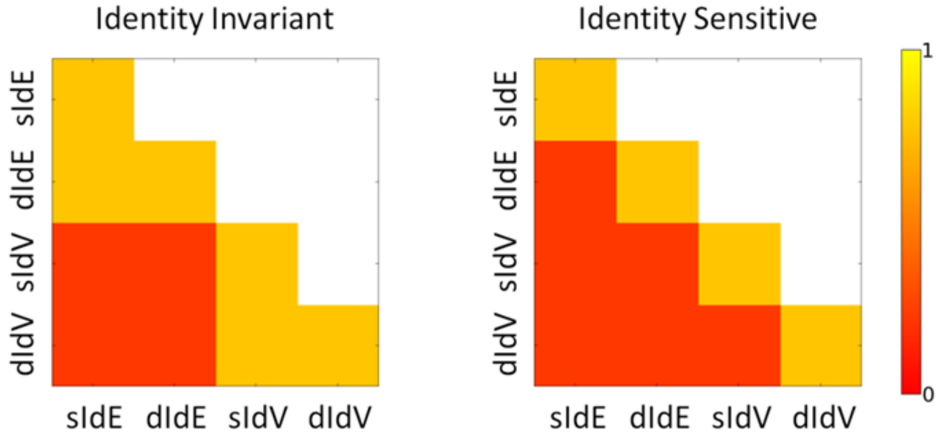


Figure 4.4. Correlation matrices showing idealised models which would represent a pattern of response that was either invariant to changes in identity (identity invariant) or sensitive to changes in identity (identity sensitive). sIdE: same identity, different expression, dIdE: different identity, different expression, sIdV: same identity, different viewpoint, dIdV: different identity, different viewpoint.

Using a simple linear regression analysis, the identity invariance model was able to explain a significant proportion of the variance in the OFA: B = 0.45, p < .001, FFA; B = 0.31, p < .001, STS; B = 0.43, p < .001 and Amgydala; B = 0.08, p < .001 (Figure 4.5). The identity sensitive model was not able to explain a significant proportion of the variance in any of the ROIs.

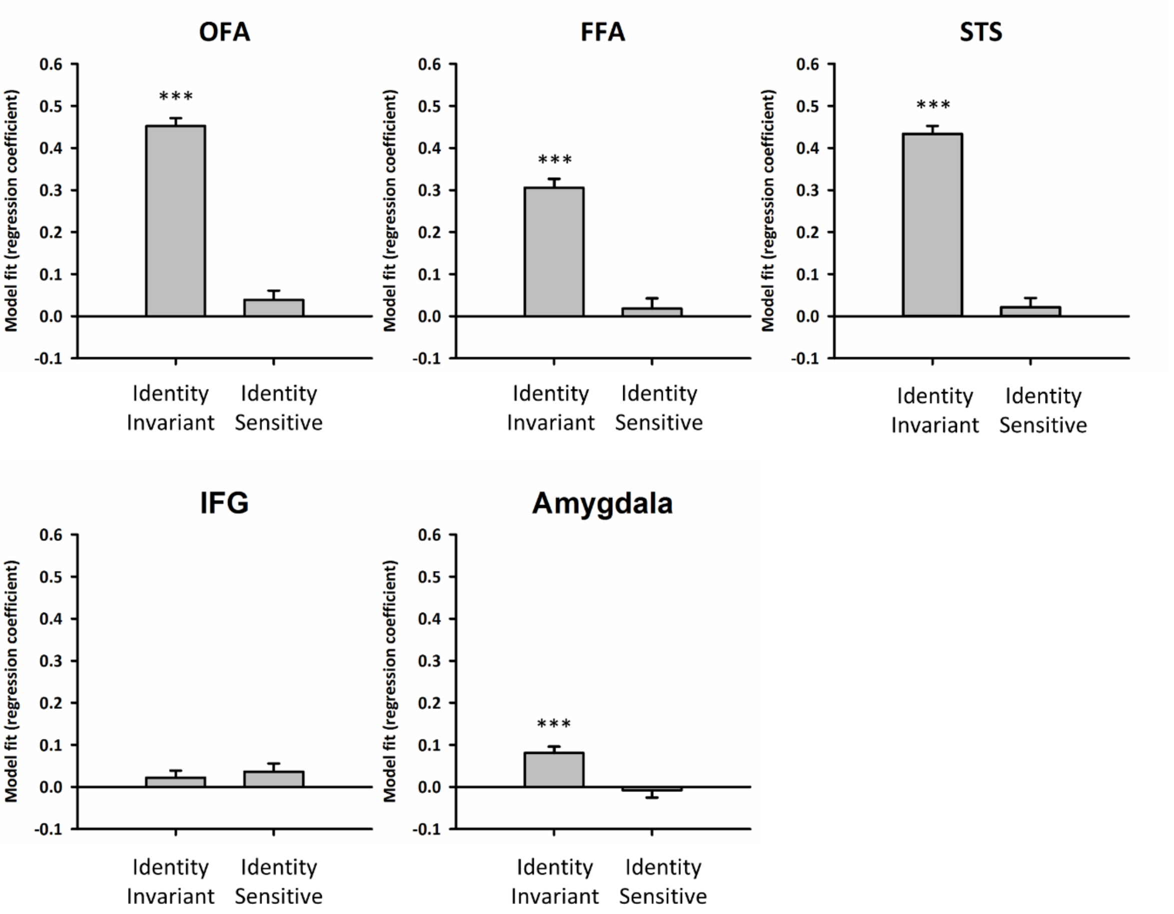


Figure 4.5. Results from the regression analysis demonstrating the higher performance of the identity invariant model in explaining the neural data.

4.4.2 Experiment 2

In Experiment 2, fMR adaptation was used to measure the selectivity of response to expression and viewpoint within the face-responsive ROIs. Release from adaptation to expression was demonstrated by voxels which responded more to different expression blocks compared to same expression blocks, irrespective of changes in viewpoint (1: dEdV > sEdV, 2: dEsV > sEsV). Release from adaptation to viewpoint was demonstrated by voxels which responded more to different viewpoint blocks compared to same viewpoint blocks, irrespective of changes in expression (3: dEdV > dEsV, 4: sEdV > sEsV). Figure 4.6 shows the percentage of voxels that released from adaptation to expression, viewpoint or both.

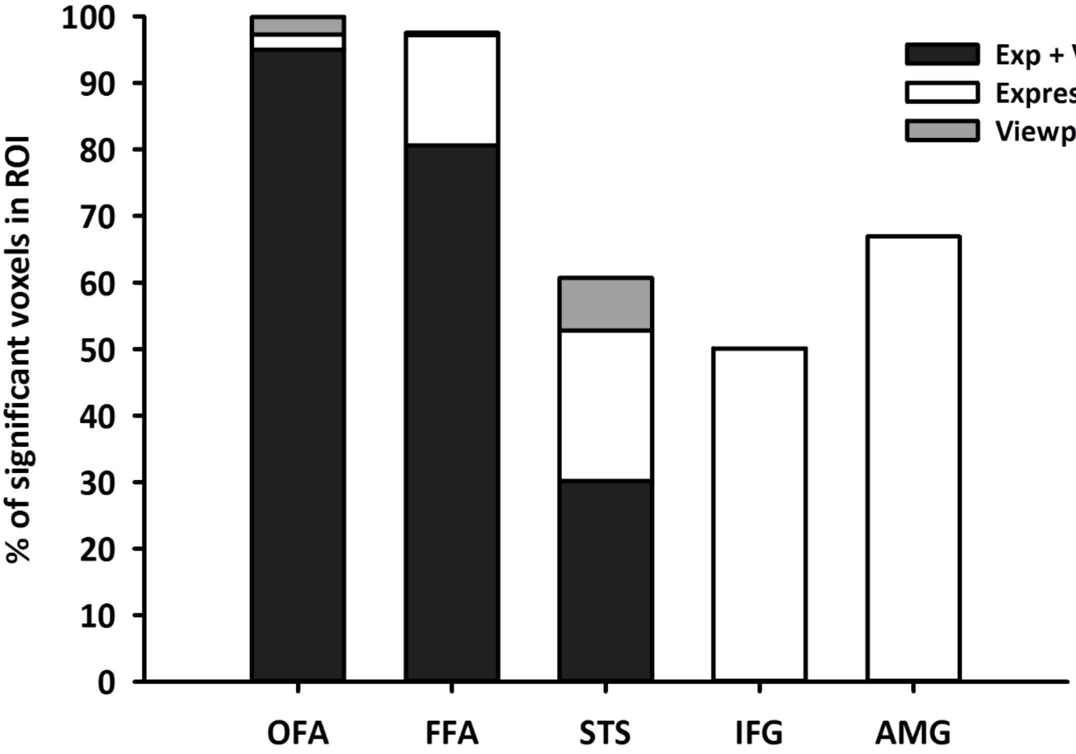


Figure 4.6. Percentage of voxels in each ROI that showed release from adaptation to expression (white), viewpoint (grey), or both (black). Release from adaptation to viewpoint was seen in the FFA, but only to 0.3% of voxels.

The majority of voxels in the OFA and FFA released from adaptation to both expression and viewpoint with little selectivity for only one category. The STS however, shows more variation with the majority of voxels releasing from adaptation to both expression and viewpoint, however 23% of voxels showed selectivity for expression, and

7.9% for viewpoint. Both the IFG and amygdala only released from adaptation to changes in expression.

4.5 Discussion

The aim of this chapter was to determine how social signals conveyed by changes in expression and viewpoint in the face are represented in face-responsive regions. Experiment 1 used MVPA to demonstrate that there are distinct patterns of response to changes in expression and viewpoint across the core face-responsive regions. The patterns of response to viewpoint and expression were consistent across individual participants, but were largely invariant to changes in facial identity. Experiment 2 used an fMR adaptation paradigm to demonstrate the level of selectivity for expression and viewpoint varies across the face-responsive regions. The OFA, FFA and STS released from adaptation to expression and viewpoint, whereas the IFG and amygdala only demonstrated adaptation for expression.

Models of face perception propose a core system in which these changeable aspects of faces are to some extent considered to be processed independently of facial identity (Bruce & Young, 2012; Haxby et al., 2000). The STS is thought to be critical for the processing of changeable aspects of faces, forming part of a core neural system for face perception that can interact with regions in an extended face processing network, such as the amygdala and parts of the frontal and parietal lobes (Allison et al., 2000). These models have often considered that the processing of changeable aspects of faces involves one core neural pathway. However, different types of change can convey very different social meanings, and involve characteristically different changes in visual information. So, it is possible that these distinct changes in the face might be represented by different patterns of neural response.

To address this issue, fMRI was used to measure the response to blocks of faces varying in expression (with a fixed viewpoint) or varying in viewpoint (with a fixed expression). The results demonstrate that there are distinct patterns of response to changes in facial expression and viewpoint in the core regions of the OFA, FFA and STS.

These representations of expression and viewpoint were largely invariant to changes in facial identity. A slightly different result was seen for the extended system of the IFG and amygdala. Although patterns of response to expression and viewpoint were less distinct from one another in these regions, a k nearest neighbour classifier was able to decode expression in both. This was supported by the adaptation results where these regions showed release from adaptation only to changes in expression and not viewpoint. The amygdala's response was also invariant to changes in identity.

This chapter found that the STS showed distinct patterns of response to changes in expression and changes in viewpoint. Consistent with these findings, Experiment 2 was able to demonstrate that different regions of the STS showed adaptation to changes in viewpoint or changes in expression. These findings are consistent with previous findings which have associated this region with the processing of changeable aspects of faces (Allison et al., 2000; Harris, et al. 2012, Psalta et al., 2014). The segregation of the neural processing for expression and viewpoint changes demonstrated in this chapter fits with differences in the information conveyed by changeable facial features (Bruce & Young, 2012). Non-rigid changes of the facial muscles can provide signals about a person's emotional state, but also provide useful cues from the movement of the lips that play an important role in understanding speech. In contrast, rigid changes of the head create upward and downward movements that can be perceived as indicating dominance or submissiveness. Additionally, lateral changes (of the head or eyes) can signal a shift in a person's focus of attention. All of these changes are considered as changeable aspects of faces in current models of face perception, but the visual information that is changing and the information that is being conveyed can be quite distinct.

The distinct patterns of response to expression and viewpoint changes may also be driven by the objective visual properties of the faces themselves. Movements of the facial musculature that accompany changes in expression lead to significant changes to the configuration of the internal features of the face. Changes in viewpoint can produce dramatic changes in the pattern of luminance across the image by exposing and occluding different regions of the face and head. So, it is also possible that these changes in image properties may have influenced the pattern of results - recent work has shown that image properties of objects can predict patterns of response in high-level visual areas (Rice et al., 2014; Watson et al., 2014). These findings suggest that the distinct patterns of response that were found may reflect the distinct image properties that are created by facial expression and viewpoint changes.

This chapter also demonstrated that the patterns of response in the OFA and FFA were able to discriminate between expression and viewpoint and that these regions also showed adaptation to both expression and viewpoint. The FFA is often considered to be part of a pathway involved in the processing of invariant features of face processing, however the FFA still performs well in discriminating expression and viewpoint, which are considered changeable (rather than invariant) features of faces. Given the role of the OFA in the early processing faces, this region’s ability to distinguish between patterns of response to changes in expression and viewpoint could support the possibility of the responses being driven by the low-level image properties. Further research is needed to explore the effect of low-level image properties on the patterns of response to changeable aspects of faces.

Information about these changeable aspects of faces is then thought to be relayed to regions in the extended face processing network for further analysis (Haxby et al., 2000). Although the patterns of response were less distinct compared to the core regions, it was possible to discriminate changes in expression and viewpoint in the amygdala and changes in expression in the IFG. In Experiment 2, adaptation for expression only was seen in these regions. The amygdala has long been implicated in the processing of facial expression, particularly that of fear and anger (Adams et al., 2003; Adolphs et al., 1995; Harris, Young, & Andrews, 2014; Mattavelli et al., 2014) and is identified by Haxby et al. (2000) as an extended region in the distributed neural system for face perception. Although not directly implicated by Haxby and colleagues, the IFG is identified by an alternative model of face perception (Ishai, 2008) and is also implicated in the processing of facial expression (Carr et al., 2003; Dapretto et al., 2006; Ishai et al., 2005).

The distinct patterns of response to different object categories revealed by MVPA have usually been thought to be based on fine-scale topographic representations that are specific to each individual (Haxby, 2012; Haynes & Rees, 2006; Tong & Pratte, 2012). Thus, in almost all human MVPA studies, the analysis is specific to each individual participant. So, an important question is the extent to which classification-based decoding strategies might generalize across individual participants (Haxby et al., 2011). In the current study a modified version of Haxby et al.'s (2001) method was used, where the pattern of response in one participant was correlated with that of the average pattern of response for the rest of the group (minus that individual) (Rice et al., 2014). Using this method allowed this chapter to demonstrate that the patterns of response to expression and viewpoint were consistent across individual participants, implying a common topographic organisation for neural mechanisms involved in face perception.

In conclusion, this chapter has demonstrated there were distinct patterns of neural response to changes in expression and changes in viewpoint in the core face-selective regions of the human brain. This finding is consistent with the distinct information conveyed by these signals, and offers an important new insight into the neural representation of changeable social signals communicated through the face. However, it is not clear whether there are there distinct representations for different facial expressions or different viewpoints. This question will be addressed in the next chapter.

Chapter 5 – Patterns of Response to Viewpoint Directions

5.1 Abstract

Viewpoint plays an important role in face perception. The human brain needs to use this information to monitor an individual's focus of attention, but also ignore these image changes for the recognition of identity. The aim of this chapter was to determine how individual viewpoint directions are represented in the human brain and if these representations differ across the core face-responsive regions. fMRI was used to investigate the underlying neural representations of viewpoint in the OFA, FFA and pSTS. Participants viewed faces that varied in direction of viewpoint (0^o. 45^o, 90^o, 135^o, and 180^o). Distinct patterns of response were found to individual viewpoints in each face region. Next, a regression analysis was used to determine whether the neural representation in each region had a head-direction (left vs right) or mirror-symmetric (left = right) pattern. Head-direction patterns of response were evident in the OFA, whereas mirror-symmetric patterns of response were found in the FFA and STS. These results suggest that viewpoint is a dominant organizing principle in the topographic organisation of face-selective regions, but that viewpoint is represented differently in different regions.

5.2 Introduction

The previous chapter addressed how changes in expression and changes in viewpoint are represented in the human brain. It was possible to demonstrate distinct patterns of response to either changes in expression or changes in viewpoint. This chapter builds on these findings by asking whether it is possible to demonstrate distinct neural patterns of response to different facial viewpoint directions.

Viewpoint plays a significant role in face recognition as it is important for us to be able to identify faces from various viewpoint directions. Changes in viewpoint are also important to determine an individual's focus of attention. Humans are remarkably good at discriminating different viewpoints with the ability to detect small differences from head outlines or internal features alone, and over a wide range of size and spatial frequencies (Wilson, Wilkinson, Lin, & Castillo, 2000). This suggests there may be an underlying neural representation for different viewpoint directions. Support for an underlying neural representation for viewpoint has been found in single neuron recordings in monkeys. Cells were found in the macaque STS that are selective for specific viewpoints of the face (Perrett et al., 1991). Indirect support for this type of cell has been found in humans using behavioural adaptation. Fang and He (2005) found that after adapting to a face with a viewpoint to the left or right, participants reported that a frontal face was perceived to be orientated in the opposite direction to that of the adapted image.

Although the processing of viewpoint direction is important for social interactions, it is also important to discount changes in viewpoint for the recognition of identity (Booth & Rolls, 1998; Kourtzi et al., 2003). Models of face processing have suggested that the recognition of facial identity is based on face recognition units that have a view-invariant representation (Bruce and Young, 1986; Burton et al., 1999). However, it is not clear how the brain achieves a view-invariant representation. It has been suggested that viewpoint-symmetric representations may represent an intermediate step between viewpoint specific coding and full viewpoint invariance (Kietzmann et al., 2012). Evidence for the representation of mirror-symmetric views has previously been demonstrated for faces in monkeys (Freiwald & Tsao, 2010; Perrett et al., 1991). Freiwald and Tsao (2010) found mirror-symmetric representations of identity in the macaque anterior lateral face patch and Perrett et al. (1991) found similar results in the STS. These studies report the existence of neurons with bimodal tuning curves, where selective responses are evident to symmetrically opposing viewpoint directions. In contrast, neuroimaging studies in humans provide mixed support for mirror-symmetric responses. Kietzmann et al. (2012) found mirror-symmetric representations for viewpoint in a large number of higher visual areas including the OFA and FFA, but not in the STS. These regions demonstrated similar neural coding for mirror-symmetric views (e.g. +90°and -90°). Axelrod and Yovel (2012) reported mirror-symmetric representations in the FFA and STS, but not in the OFA. In contrast, Ramirez, Cichy, Allefeld and Haynes (2014) reported a strong frontal viewpoint representation in the FFA, but could not find reliable evidence of a mirror-symmetric representation.

Together, previous studies suggest that the neural processing of faces requires different representations of viewpoint. Using EEG, Kietzmann, Gert and König (2015) found a distinct temporal sequence of coding schemes beginning with an early stage assessing head orientation, followed by viewpoint symmetric representations. However at a later stage, viewpoint invariance was seen but crucially this was excluding the face-on view. These assessments of viewpoint direction have various potential uses. They suggest head orientation coding is important for shared visual attention. This can be particularly useful from an immediate assessment of the situation (what is going on and where), but also from a social communication point of view (what is this person looking at and why). Mirror symmetric coding of viewpoint direction is potentially an important step towards full viewpoint-invariance for identity recognition.

The aim of this chapter was to investigate the neural coding behind different viewpoint directions in the face-selective regions of the OFA, FFA and STS. Given the proposed roles of the different core regions (Haxby et al., 2000), it is possible that these regions may demonstrate different representations of viewpoint directions. The FFA is associated with the processing of invariant aspects of faces such as identity, so is likely to have a representation that is invariant to changes in viewpoint, such as symmetric coding. The STS is associated with processing changeable aspects of faces that are important for social communication, so it is possible that this region only has distinct representations for different directions of viewpoint. The OFA is an early stage of processing so this region may have a view-dependent pattern of response.

5.3 Method

5.3.1 Participants

There were 20 right-handed participants (10 female, mean age 25.3 ± 3.1). All participants had normal or corrected-to-normal vision with no history of neurological conditions. Written consent was obtained from all participants and the studies were approved by the York Neuroimaging Centre Research Ethics Committee.

5.3.2 Stimuli and Design

An independent localiser scan was used to define group level regions of interest, details of which can be found in Chapter 2, section 1.2.2. *Functional localisation at the group level.* Accuracy on the red dot task was high (98.6 + 5.0%), indicating that participants were paying attention to the stimuli.

In the main experiment, there were seven stimulus conditions: 1) *right profile (*RP*)*, 2) *right ¾ profile* (R¾), 3) *front view* (FR), 4) *left ¾ profile* (L ¾), 5) *left profile* (LP), 6) *mixed viewpoints,* and 7) *Fourier scrambled versions of conditions 1-5.* Conditions 6 and 7 were intended as an inbuilt localiser, however were not used in the following analyses. Each condition contained black and white images from five different identities posing a neutral expression. Examples of the images from the first five conditions are shown in Figure 1. Face images were taken from the Radboud Faces Database (Langner et al., 2010) and placed onto a 1/f amplitude mask. This was to ensure that all images stimulated the same amount of the visual field despite changes in orientation. Images were presented approximately 57cm from the participant and were approximately 9° x 8°. Images from each condition were presented in 6 second blocks, with each block containing 5 images. Each image was presented for 1000 ms with a 200ms grey screen ISI. There was a 9 second inter-block period during which a grey fixation screen was presented. Each stimulus condition was repeated 6 times. There were 42 blocks in total. Participants monitored images for the presence of a red dot. Accuracy on the task was high, with a mean accuracy of 99.9% (SD 0.5) and mean response time of 454 ms.

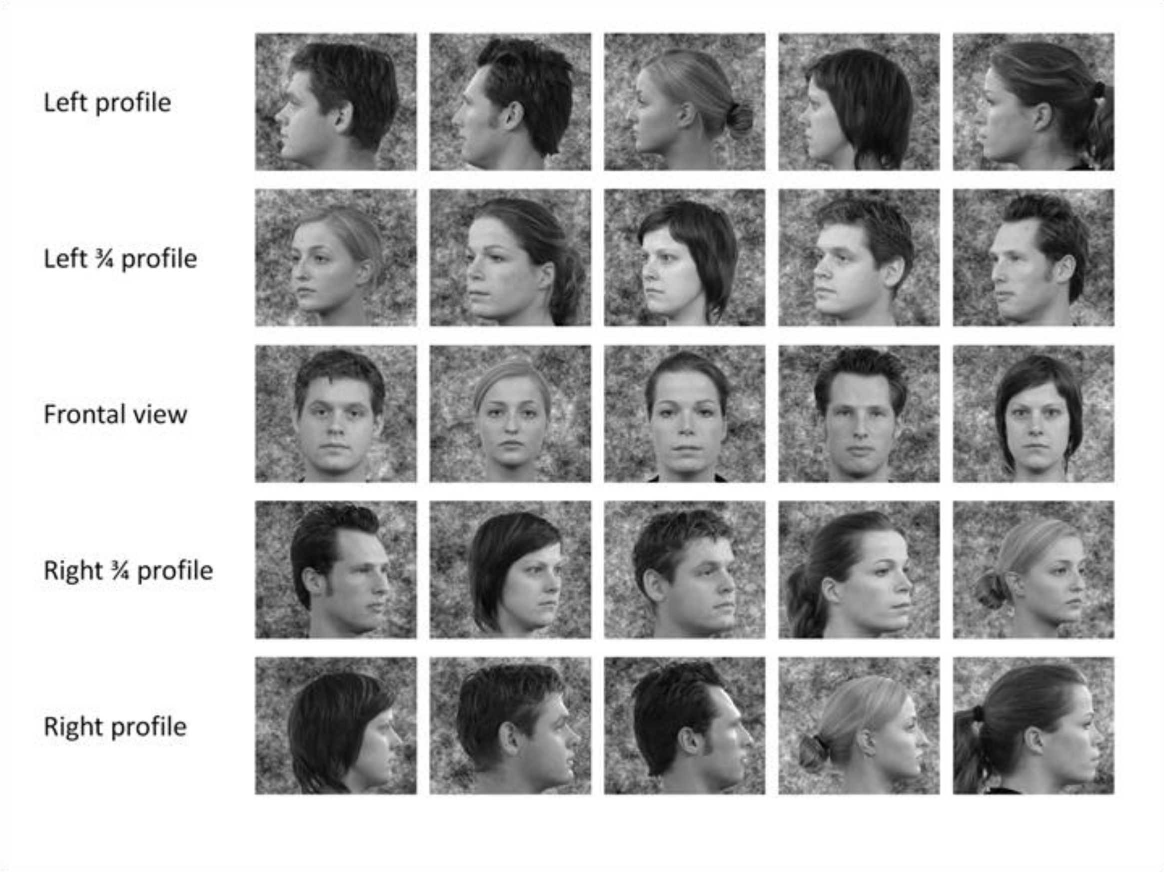


Figure 5.1. Examples of stimuli from the main experimental conditions. Each row shows the sequence of images in a representative stimulus block from the different conditions.

5.3.3 Imaging Parameters

All scans were conducted using a 3 Tesla MRI system with an eight channel phased array head coil (GE Signa Excite HDx 3.0T, High resolution brain array, MRI Devices Corp., Gainesville, FL) tuned to 127.4MHz. A gradient echo planar imaging (EPI) sequence was used to acquire the data. The acquisition parameters were: 38 contiguous axial slices, repetition time (TR) 3 seconds, echo time (TE) 32.5 milliseconds, flip angle 90°, field of view (FOV) 28.8 x 28.8 cm, matrix 128 x 128, slice-thickness 3mm, voxel size 2.25 x 2.25 x 3mm. To improve registration, the EPI image was co-registered with a T1-weighted image taken in the same plane, before being registered to the high resolution main structural scan (T1-weighted, 1.13 x 1.13 x 1 mm) of each participant. This was then co-registered to the standard MNI 152 brain.

5.3.4 fMRI Analysis

ROIs for the subsequent MVPA analysis were defined as described in Chapter 4, section 4.3.4.1 *Experiment* 1. Pattern analyses were performed as described in Chapter 2, section 1.3.2.3 *LOPO* Methods. This LOPO method was used to determine the consistency of the patterns of response across participants by measuring how similar each participant's responses were to those for the rest of the group.

To assess whether there are distinct patterns of response to individual viewpoint directions, a repeated measures ANOVA was used to test the difference between the average within-condition (e.g. RP-RP, FR-FR) and between-condition (e.g. RP-LP, FR-LP) correlations for each LOPO iteration across the three ROIs. If a direction evoked a distinct pattern of response, then the within-condition correlations for the individual participant and rest of the group data should be higher than the between-condition correlations in the given region.

A regression analysis was then used to assess whether there are differences in the way that viewpoint is represented. For each factor, a binary regressor was generated representing an idealised correlations matrix in which 0 or 1 was entered in different cells. Two regressors were created which represented patterns of response which would suggest a representation for left or right (direction), or partial view-invariance representation (symmetry). Each regressor was then entered into a simple linear regression, with the outcomes defined as the correlations matrices obtained from the MVPA concatenated across LOPO iterations. A direction and symmetry representation would both expect higher within-category correlations. As this may drive a high correlation between the matrices regardless of underlying representational structure, the analysis was performed without the within-category correlations (the on-diagonal cells). This analysis yielded a beta value for each regressor which would be expected to be significantly greater than zero if that regressor was able to explain a significant amount of the variance in the MVPA data.

Next this study addressed whether the patterns of response could be explained by the image statistics of the facial viewpoints. Image statistics of each viewpoint direction were computed using the GIST descriptor (http://people.csail.mit.edu/torralba/code/spatialenvelope/). For each image, a vector of 512 values was obtained by passing the image through a series of Gabor filters across eight orientations and four spatial frequencies, and windowing the filtered images along an 8 x 8 grid. Each vector represents the image in terms of the spatial frequencies and orientations present at the different positions across the image. A cross-validation procedure was used to determine how similar individual face viewpoint images were to the average of each viewpoint direction. GIST descriptors were averaged across all but one of the images within each viewpoint direction. These average descriptors were then compared with each unique image creating within- and between-viewpoint correlations for each combination of viewpoint directions. The correlation values for the GIST descriptor across different viewpoint directions were represented in a correlation matrix which was then used as a regressor for the fMRI data in the same manner as described above, again removing the on-diagonal cells as both the GIST matrix and fMRI data have a strong on-diagonal component.

5.4 Results

Figure 5.2A shows the similarity of the patterns of response to different facial viewpoints across the core face-selective regions. The first aim was to determine whether there were distinct patterns of response to each viewpoint. To do this, within-viewpoint correlations (diagonal) were compared to between-viewpoint correlations (off-diagonal). The results were entered into a 3x2 repeated measures ANOVA, with the factors region (OFA, FFA, STS) and condition (within, between). The results demonstrated a significant main effect of condition, with higher within-viewpoint correlations, compared to between-viewpoint correlations ($F(1,19) = 35.18$, $p < .001$). This suggests distinct representations for individual viewpoint directions in the core face-responsive regions. There was no significant main effect of region or an interaction between the main effects. The results show that higher correlations were seen along diagonal elements, when compared with off-diagonal elements, suggesting distinct representations for individual viewpoint directions (OFA: $t = 4.44$, $p < .001$, FFA: $t = 4.28$, $p < .001$, STS: $t = 4.34$, $p < .001$).

The next step was to compare the neural representation behind different viewpoint directions in the core face-selective areas as identified by Haxby et al., (2000). Specifically this chapter addressed whether there are patterns of response that suggest head orientation encoding and partial view-invariance (mirror-symmetric representations). MVPA was used to investigate the similarity of the patterns of response to the five different viewpoint conditions across participants, 1) *right profile (*RP*)*, 2) *right ¾ profile* (R¾), 3) *front view* (FR), 4) *left ¾ profile* (L ¾), 5) *left profile* (LP).

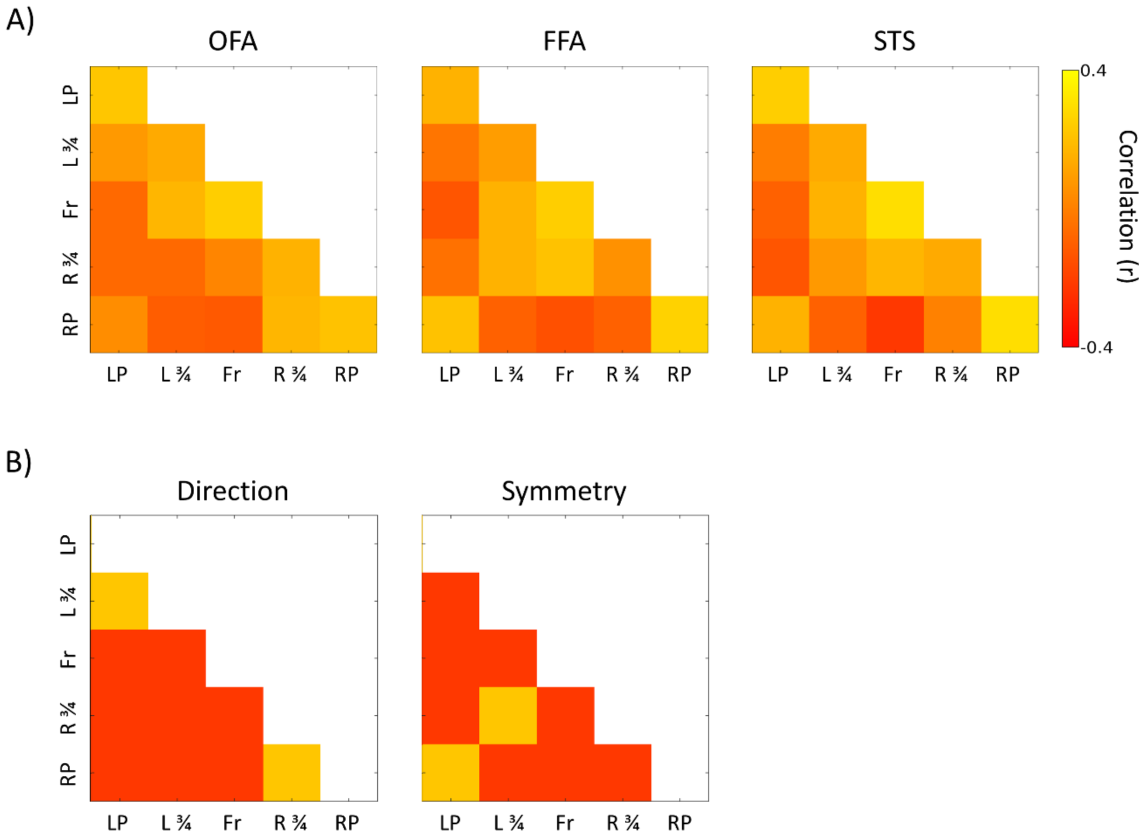


Figure 5.2. Correlation matrices showing similarity in the patterns of response between conditions in each of the core face-selective regions (A). RP: right profile, R ¾ : right ¾ profile, Fr: frontal view, L ¾ : left ¾ profile, LP: left profile. (B) Model predictions for head orientation (direction – left/right) and partial view-invariance (symmetry). These models were entered into the regression analysis.

To address whether there may be different representations of viewpoint directions, across the face-responsive areas, models were created based on prior research discussed above (Kietzmann et al., 2015). These regressors were generated to represent an idealised coding of head direction and symmetry (Fig. 2B). In the OFA the direction selective model was able to predict the variance in the neural response (B = 0.25, p < .001). In contrast, the symmetry model was not able to explain the neural response (B = -0.05, p = 0.323). In the FFA, the symmetry model was able to explain a significant proportion of the variance (B = 0.27, p < .001), whereas the direction model negatively predicted the response profile in the FFA (B = -0.14, p < .010). In the STS the symmetry model explained a significant proportion of the variance (B = 0.23, p < .001), but the direction model was not able to predict the neural data (B = -0.01, p = 0.882). The regression analysis therefore demonstrates a differing response profile across the three regions (see Figure 5.3). The OFA shows a greater model fit from the direction selective model, whereas the FFA and STS are more consistent with the symmetry model.

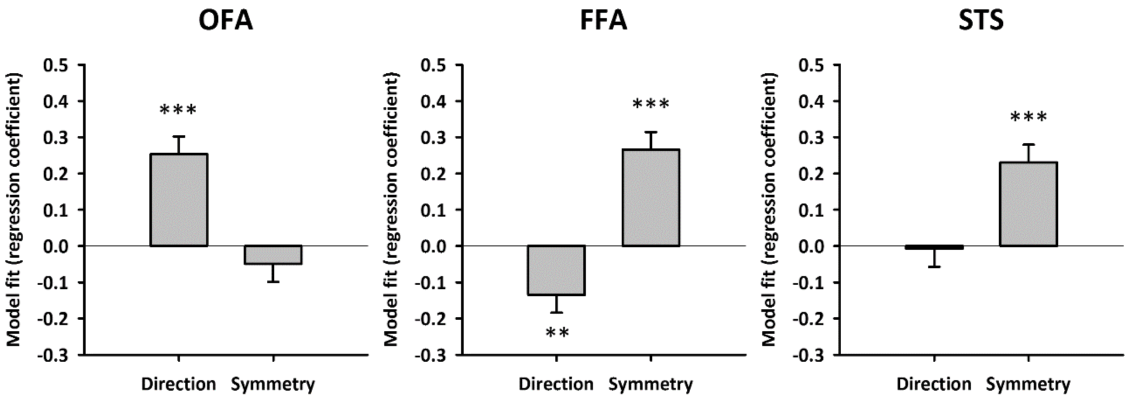


Figure 5.3. Results from the regression analysis demonstrating the different response profiles across the face-selective ROIs

Next this study addressed whether the response profiles in the face-responsive ROIs could be explained by differences in the low-level image statistics for each of the viewpoint directions. The image statistics of each viewpoint direction were analysed using a GIST descriptor (Oliva & Torralba, 2001). Figure 5.4 shows the correlations in image properties within-viewpoint and between-viewpoint directions.

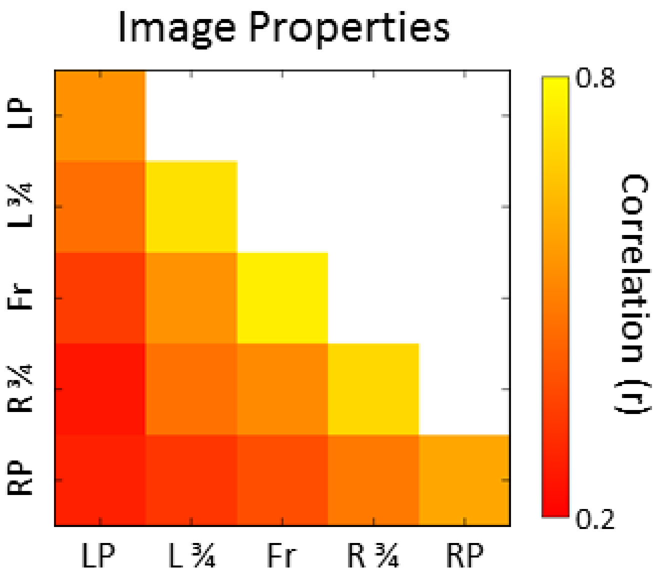


Figure 5.4. Correlation matrix demonstrating the similarity in image properties within- and between-viewpoint direction. A clear diagonal can be seen, where the within-viewpoint correlations are high, particularly those of the frontal views and the viewpoints either side. Lower similarity is seen between the image properties of opposing viewpoint directions, e.g. RP vs LP.

The GIST descriptor matrix (minus the on-diagonal elements) was used as a regressor for the fMRI MVPA matrices, in a simple linear regression to determine how much of the variance in the patterns of response could be explained by the orientation and spatial frequency information in the images. The GIST descriptor was able to explain a significant proportion of the variance in all regions (OFA: B = 0.26, $p < .001$, FFA: B = 0.15, $p < .010$, STS: B = 0.24, $p < .001$).

5.5 Discussion

The aim of this chapter was to determine how facial viewpoint directions are represented in core face-responsive regions of the human brain. Using MVPA, distinct patterns of response to different viewpoints were found in each region. However, there were differences in the way that viewpoint was represented in each region. A more mirror-symmetric representation of viewpoint was found in the FFA and STS. In contrast, the neural representation of the OFA was more tuned to overall head direction.

Models of face perception suggest different roles for the core regions involved in face processing (Haxby et al., 2000; Ishai, 2008). The OFA is proposed to be involved in the early perception of faces, the FFA is identified as processing invariant aspects of faces such as facial identity, and the STS is thought to be important for processing the changeable aspects of faces. Therefore, this study addressed the response to facial viewpoint in these three core regions and to see if facial viewpoint may have differential representations across these three regions.

The results demonstrated evidence of viewpoint-specific coding in the core regions of the OFA, FFA and STS. This supports previous work in macaques which found cells with unimodal responses to a specific viewpoint direction (Freiwald & Tsao, 2010; Perrett et al., 1991). Fang and He (2005) used adaptation in humans to demonstrate evidence of neural populations in the visual system that have specific viewpoint direction coding and this was later confirmed using fMRI by Axelrod and Yovel (2012) who were able to decode facial viewpoint in the OFA, FFA and STS.

Based on previous literature, different models of viewpoint representation were developed which could be used to probe the neural representation in each face region. The neural response in the OFA best reflected the head-direction model. This model represented an idealised response to orientation selectivity (to the left or to the right). This result is consistent with the early EEG response patterns to head orientation (Kietzmann et al., 2015). Given the early nature of this response, < 100 ms after stimulus onset, it would be expected that this response would be associated with regions involved in the initial processing of facial viewpoint, such as the OFA. The OFA is considered to be the first stage of processing in models of face perception (Bruce and Young, 1986; Haxby et al., 2000; Ishai, 2008) and its causal role in this time frame is supported by TMS work showing disruption in face perception when TMS pulses are applied to the OFA at around 60-100ms (Pitcher, Garrido, Walsh, & Duchaine, 2008; Pitcher et al., 2007).

A more mirror-symmetric neural representation was found in the FFA and STS. That is, more symmetric viewpoints (right profile – left profile) were more similar than non-symmetric viewpoints (right profile – left ¾). These findings are consistent with the development of a viewpoint specific representation in early visual areas, to a partial view-invariant representation (invariance for mirror-symmetric profiles), and finally to a full view-invariant representation in higher order visual areas (Axelrod & Yovel, 2013; Kietzmann et al., 2012). To some extent, this fits with the Haxby et al. (2000) neural model, where the FFA is implicated in the processing of facial identity. This would suggest a degree of viewpoint invariance for identity. The STS however is implicated in the processing the changeable aspects of faces, and therefore might be expected to have a viewpoint specific representation, in order to process social relevant cues such as focus of attention.

The dissociation seen in this study between the OFA and the FFA and STS, is similar to that found by Axelrod and Yovel, where they found mirror-symmetric representations in the FFA and STS but not in the OFA. Decoding accuracy of the different views was higher in the OFA and this is again in line with the findings of this chapter, where the data is better explained in the OFA by the direction model, and higher within viewpoint correlations compared to between viewpoint correlations. Freiwald and Tsao (2010) also demonstrated similar results in macaque monkeys where they found view-selectivity in macaque posterior face patches, but more mirror-symmetric coding when moving to middle face patches and finally full invariance in the most anterior regions.

The ability of the head direction model to explain the data in the OFA and the symmetry model in the FFA and STS, dovetail nicely with previous research into the temporal dynamics of facial viewpoint encoding (Kietzmann, Gert, & König, 2015). Using EEG, Kietzmann et al. demonstrated that head orientation was decoded initially, with similar views leading to similar responses, reflecting the representation found in the OFA in this chapter. Shortly after, effects of viewpoint symmetry were seen in the EEG response patterns, fitting nicely with the symmetry representation found in the present study, for the FFA and STS. At a later stage, Kietzmann et al. found the response patterns demonstrated near viewpoint invariance, with the only view showing selectivity being the frontal view. Selectivity for the frontal view could be important for regions involved in social interaction, as tracking small changes in eye gaze and expressions can give a wealth of information to the observer and aid social communication. As yet, there are no studies that have found viewpoint invariance, or near viewpoint invariance in specific neural regions in humans, however this was demonstrated in macaques by Freiwald and Tsao (2010).

One reason why different viewpoints might elicit different patterns of response is that the image properties of faces from different viewpoints varies considerably. To address this issue, a low-level image descriptor was used to compare the similarity of faces at different viewpoints. Interestingly, the patterns of response to facial viewpoint in the core face-selective regions could be explained by the similarity in low-level image statistics. This is consistent with previous studies suggesting patterns of response in category-selective regions can be predicted by low-level image properties (Rice et al., 2014). However, this cannot explain the similarity in neural responses for the mirror-symmetrical viewpoint directions.

It is important to note that although 'mirror-symmetric view' is used to describe the face images which represent left and right profiles or ¾ profiles, this study used images from the Radboud database (Langner et al., 2010) of real human faces simultaneously captured from a number of viewpoints. So, the images were not truly mirror symmetric. Human faces are less symmetrical than might initially be thought and research suggests that the right side of the face is more diagnostic of the whole face, although this is likely to be due to the right hemisphere dominance in face perception (Gilbert & Bakan, 1973). Given this, it is important to use true human faces to study viewpoint perception, as if the brain uses symmetry it needs to be able to account for the variability in the left and right sides of the face. However it would be interesting in the future to compare the results between true mirror-symmetric views, and real symmetric views.

An important point to consider is that in this experiment unfamiliar faces were used. This raises the question of whether the neural coding of viewpoint changes would be different for familiar faces. Ewbank and Andrews (2008) suggest sensitivity to viewpoint differs between familiar and unfamiliar faces in the FFA. Adaptation was seen to familiar faces across a range of viewpoints, whereas release from adaptation was seen for unfamiliar faces. This opens the possibility for future work to determine whether the patterns of response to different facial viewpoints for familiar faces, might be represented differently across core face-selective regions.

In conclusion, a mirror-symmetric representation of profile and ¾ profile face images was seen in the face-selective areas of the OFA, FFA and pSTS. A representation consistent with direction-selectivity was seen for the OFA, whereas partial view-invariance was supported in the FFA and pSTS. This chapter addressed the representation of viewpoint only, and so this leads to the question of how the representation of specific viewpoint directions might be modulated by facial expressions (and vice versa) in the next chapter.

Chapter 6 – Patterns of Response to Specific Viewpoints and Expressions in Face-Selective Regions of the Human Brain

6.1 Abstract

Models of face processing suggest that changeable aspects of faces, such as viewpoint and expression are processed in the same regions. Chapter 4 found distinct neural representations of changes in expression and changes in viewpoint. In this chapter, the relative role of viewpoint and expression in the topographic organization of face regions was compared directly. fMRI was used to measure the neural response to different combinations of facial expression and viewpoint combinations. Participants (n = 24) viewed blocks of faces that contained one of three expressions (happy, fear, disgust) and one of three viewpoint directions (left ¾ profile, frontal view, right ¾ profile). Using MVPA, distinct patterns of response for each of the viewpoint directions were found in the OFA, FFA and STS. However these distinct representations for the individual viewpoints, were independent of the facial expression posed, suggesting that viewpoint is the dominant feature. To determine whether more basic dimensions of the stimulus could explain these patterns of neural response, the image properties were measured. A strong positive correlation between the neural patterns and the underlying low-level image properties was evident in the core face regions. These results suggest that the neural representation of faces may be based on the statistical properties of the image rather than higher level attributes, such as viewpoint and expression.

6.2 Introduction

Facial movements are important for effective social communication. These changes can also convey information which is important for survival. As discussed previously, movements of the face have traditionally been split into two categories; rigid and non-rigid movements (Bruce & Young, 2012). In the previous studies presented in this thesis, distinct patterns of response were found to changes in viewpoint (rigid movement) and changes in expression (non-rigid movement). However, the relative influence of viewpoint on the perception of specific facial expressions has not been addressed. This study aimed to assess to what extent the neural representation of expression is influenced by viewpoint, and vice versa.

There is some evidence to suggest that facial expressions can be decoded in face-responsive regions such as the pSTS. Said, Moore, Engell and Haxby (2010) used dynamic images of facial expressions (anger, disgust, fear, sadness, happiness and surprise). Using MVPA, they were able to decode the individual facial expressions in both the posterior and anterior STS. They were also able to demonstrate a level of similarity between behavioural perceptual similarity ratings of facial expressions, with the MVPA similarity matrix. However, Zhang et al. (2016) used support vector machine pattern classification analysis (SVM) to address the neural coding of four facial expressions (fear, anger, happiness and neutral) in various face-responsive regions. Whilst the amygdala was able to discriminate fearful faces from other faces, the pSTS was only able to discriminate neutral from emotional (i.e. neutral vs. fear, anger and happiness). In addition, Srinivasan, Golomb and Martinez (2016) used MVPA to demonstrate decoding of facial action units in the pSTS, however the decoding ability of facial expressions was much less consistent and reliable. The different methods used in these studies makes it difficult to make a direct comparison, as does the type of stimuli used – dynamic vs. static. However there is reasonable evidence that it is possible to discriminate facial expressions in the pSTS.

As discussed in Chapter 5, there is also evidence to demonstrate that facial viewpoint directions can be decoded. Axelrod and Yovel (2012) were able to decode individual viewpoint directions in the OFA, FFA and STS. In addition, Kietzmann, Swisher, Konig and Tong (2012) developed a model of low-level visual similarity between five viewpoint directions. This model predicted repeated images of the same viewpoints would show a similar pattern of response, with moderate similarity between neighbouring viewpoint directions. This model was able to explain a large proportion of the variance in low-level visual areas such as V1, V2, V3 and hV4, however it was also able to explain a significant proportion of the variance in the patterns of response in all the higher order visual areas they tested, including OFA, FFA and LO. Natu et al. (2010) were also able to demonstrate above chance discrimination of viewpoint pairs. Therefore, there is a good degree of evidence of the ability to decode individual viewpoint directions and facial expressions. However, the ability to discriminate specific combinations of expressions and viewpoints has not been addressed.

The processing of viewpoint and expression are often considered to be independent; the perception of viewpoint allows us to assess the direction of attention of an individual, whereas the perception of facial expression allows us to assess how they may be feeling. However, it is also important for us to know whether what they are attending to is relevant to us. If we see that someone is fearful, it is probably not enough to know this, we also need to know where this emotion is directed. This information is crucial in situations where we may be at threat and unaware of the danger which another person perceives.

There is a large body of behavioural literature looking at the modulation of facial expressions by non-rigid movements of the head and eyes. Work by Adams and Kleck (2003) demonstrated that direct eye gaze facilitates processing of approach orientated expressions such as anger and happiness, whereas direct gaze facilities the processing of avoidance orientated emotions (e.g. fear and sadness). Hess, Adams and Kleck (2007) argue that this suggests that not only does looking direction affect the perception of facial expressions, but also the reaction to such faces in terms of formulating the most appropriate emotional and behavioural response. To this end, Hess et al. (2007) manipulated the focus of visual attention conveyed through head direction to assess how this modulates the perception of approach and avoidance orientated expressions. They found that viewpoint direction has a strong influence on the reaction to threat related expressions (fear and anger) where front facing anger expressions are more accurately decoded and produce higher levels of negative affect and anxiousness in the observer. Fearful faces with an averted head direction produced higher levels of negative affect in the observer, however they did not find averted fear was significantly better recognised than direct fear. Happiness and sadness were equally well recognised from either head direction, and the interpretation of happiness expressions was not affected by head direction. This supports the idea that specific combinations of expressions and viewpoint are more pertinent for survival, whereas non-threatening expressions such as happiness do not convey this kind of message. Hess et al. (2007) suggest that a happiness expression suggests that everything is well in the environment, and so the need to locate the area of the environment where everything is well, is less relevant.

Complementary neural results have also been found where brain regions such as the amygdala have been implicated in the modulation of facial expressions by rigid movements of the head and eyes. The amygdala is widely believed to be responsive to threat-related emotional expressions such as fear and anger (Mattavelli et al., 2014). Eye-gaze gives important information about where the fear or anger may be directed or what the source of the threat may be. As discussed in Chapter 1, there have been a number of studies looking at how rigid movements of the eyes modulate the response of the amygdala. Adams, Gordon, Baird, Ambady and Kleck (2003) used fMRI to demonstrate the amygdala's response could be modulated by direct and averted gaze in the facial expressions of fear and anger. Although there were no differences in the response to fear and anger as a function of gaze in the right amygdala, they did find significantly higher responses in the left amygdala to displays of ambiguous threat over clear threat (fear with direct gaze vs fear with averted gaze).

The research above supports the role of rigid movements of the face modulating the response to facial expressions. However this involves the use of fMRI univariate analyses, and evidence of modulation outside the region of the amygdala is limited. This study asked whether, using MVPA, it is possible to find distinct patterns of response to specific viewpoint and expression combinations. In addition, this chapter also asked whether expression or viewpoint categories might be the dominant organising principle in the core face-responsive regions. Based on the research discussed above, distinct patterns of response are likely to be seen in the amygdala, where specific combinations of expressions and viewpoints have specific meanings. Viewpoint dominant responses are more likely in regions considered to be involved in the early visual analysis of faces, e.g. the OFA, as this is a salient visual change.

6.3 Method

6.3.1 Participants and Stimuli

In Experiment 1, there were 20 participants (6 male, mean age 28.2 ± 3.3). There were 20 participants in Experiment 2 (6 male, mean age 24.2 ± 3.7). In the fMRI experiment (experiment 3) there were 25 participants (9 male, mean age 23.5 ± 2.5). One participant was excluded from Experiment 3 for excessive movement in the scanner. All participants had normal or corrected-to-normal vision. Written consent was obtained from all participants and the studies were approved by the Department of Psychology Ethics Committee at the University of York (Exp. 1 and 2) and the York Neuroimaging Centre Research Ethics Committee (Exp. 3).

Stimuli consisted of greyscale images of faces posing 3 facial expressions (fear, happiness, disgust) and had a facial viewpoint direction of either left ¾ profile, front view or right ¾ profile (Figure 6.1). The number of expressions and viewpoints were cut down from five to three each in order to reduce the number of conditions. A 5 x 5 design would have required 25 conditions, and a considerably longer experiment. The three expressions of fear, happiness and disgust were chosen so as to represent one top recognisable expression (fear), one bottom recognisable expression (happiness) and one expression equally recognisable from the top and bottom (disgust). The three viewpoints chosen were to include the front viewpoint, in addition to two symmetrical profile views. The choice between ¾ profile views rather than full profile views was to allow comparison with another study, however the results are likely to be complementary across ¾ and full profile views.

Images of 5 identities were taken from the Radboud Faces Database (Langner et al., 2010). These images were placed onto a 1/f amplitude mask to ensure all images stimulated the same amount of the visual field despite changes in orientation. Images were presented on an LCD monitor, approximately 57cm from the participant and were approximately 9° x 8°. Stimuli were presented using PsychoPy2 (Peirce, 2007).

6.3.2 Experiment 1

6.3.2.1 Stimuli and Design

In this experiment, participants rated the perceptual similarity of images which contained combinations of expressions and viewpoints. Images were presented in pairs, with the identity across the two images always being different to prevent any confounds from identity. All possible combinations of expression and viewpoint pairings were used, including the pairings containing the same expressions and same viewpoints. This was repeated for all identity combinations, totalling 450 trials. Images were presented sequentially, with each image being presented for 1000ms with a 200ms ISI between the two images. The presentation order of trials was randomised for each individual participant. Participants were required to respond with a button press indicating how similar they perceived the expressions to be, on a scale of 1 – 7 (1 being less similar and 7 being more similar).

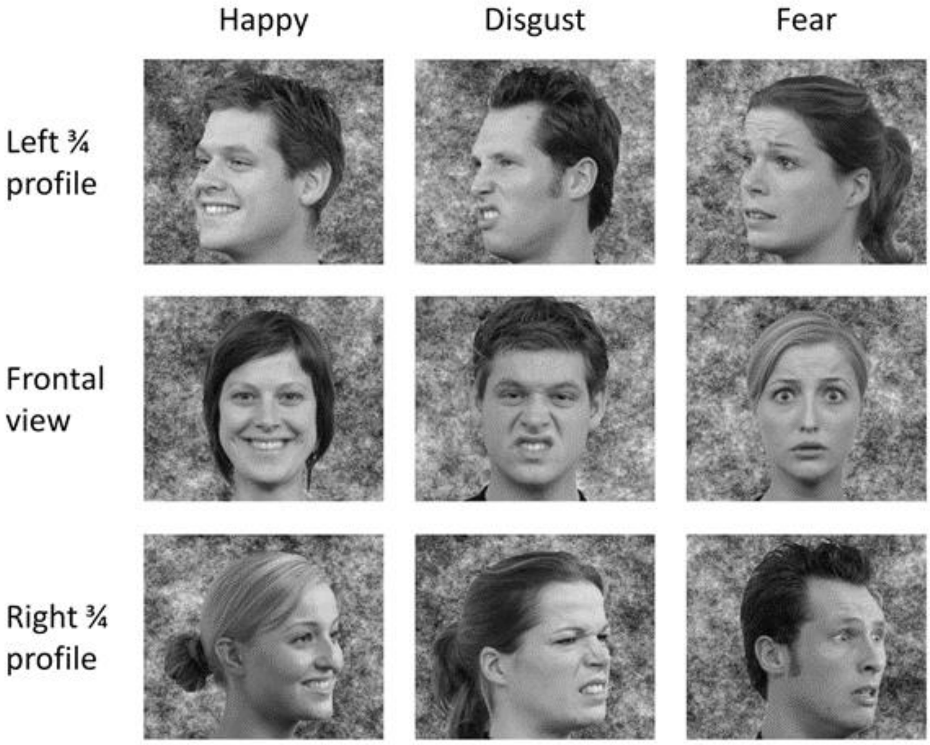


Figure 6.1. Examples of the stimuli used in Experiments 1-3.

6.3.3 Experiment 2

6.3.3.1 Design

For Experiment 2, participants again rated the perceptual similarity of images which contained combinations of expressions and viewpoints. In this experiment, participants were asked to rate the similarity of the viewpoint. All other methodological details are the same as Experiment 1.

6.3.4 Experiment 3

6.3.4.1 Design

There were nine stimulus conditions: 1) happy; left ¾ profile (HL); 2) happy frontal view (HF), 3) happy; right ¾ profile (HR), 4) fear; left ¾ profile (FL); 5) fear frontal view (FF), 6) fear; right ¾ profile (FR), 7) disgust; left ¾ profile (DL); 8) disgust frontal view (DF), 9) disgust; right ¾ profile (DR). Each condition contained images of five different identities posing the appropriate facial expression and facing in the relevant viewpoint direction. Face images were taken from the Radboud Faces Database (Langner et al., 2010) and placed onto a 1/f amplitude mask to ensure all images stimulated the same amount of the visual field despite changes in orientation.

Images from each condition were presented in 6 second blocks, with each block containing 5 images. Each image was presented for 1000 ms with a 200ms grey screen ISI. There was a 9 second inter-block interval where a grey screen with a white fixation cross was presented. Each condition was repeated 6 times, giving 54 blocks in total. To ensure participants were paying attention throughout the experiment, they were asked to monitor images for the presence of a red dot which was superimposed on one image per block, at a random location on the face. Accuracy was high, with a mean accuracy of 99.0% ± 1.9 and mean response time of 486.9 ms ± 89.7.

6.3.4.2 Imaging Parameters

All scans were conducted using a 3 Tesla MRI system with an eight channel phased array head coil (GE Signa Excite HDx 3.0T, High resolution brain array, MRI Devices Corp., Gainesville, FL) tuned to 127.4MHz. A gradient echo planar imaging (EPI) sequence was used to acquire the data. The acquisition parameters were: 38 contiguous axial slices, repetition time (TR) 3 seconds, echo time (TE) 32.5 milliseconds, flip angle 90°, field of view (FOV) 28.8 x 28.8 cm, matrix 128 x 128, slice-thickness 3mm, voxel size 2.25 x 2.25 x 3mm. To improve registration, the EPI image was co-registered with a T1-weighted image taken in the same plane, before being registered to the high resolution main structural scan (T1-weighted, 1.13 x 1.13 x 1 mm) of each participant. Linear affine transformations were calculated to align session data to intermediate, high resolution and standard (MNI) anatomical spaces using FLIRT (Jenkinson & Smith, 2001).

6.3.4.3 fMRI Analysis

ROIs were defined as described in Chapter 4, section 4.3.4.1 *Experiment 1.* An MVPA LOPO analysis was conducted on the data as described in Chapter 2, section 1.3.2.3. *LOPO Methods*.

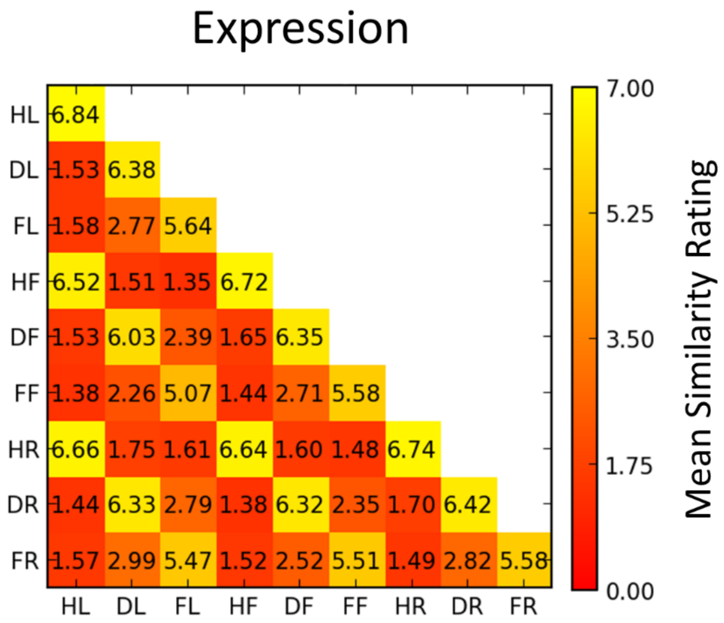
Next, a representational similarity analysis was used to explore the patterns of response. The average perceptual similarity ratings from Experiment 1 and 2 were used to produce models of expression (Exp. 1) or viewpoint (Exp. 2). The first step was to correlate the similarity matrices for expression and viewpoint with the average MVPA similarity matrix for each ROI, using Pearson's r correlation coefficient. Next, the model matrices were used as regressors, which were then entered into a simple linear regression, with the outcomes defined as the correlations matrices obtained from the MVPA concatenated across LOPO iterations. This analysis yielded a beta value for each regressor which would be expected to be significantly greater than zero if that regressor was able to explain a significant amount of the variance in the MVPA data.

This study then explored the relationship between the neural patterns of response and the image statistics of the stimuli. Image statistics of the nine conditions were computed using the GIST descriptor (Oliva & Torralba, 2001). For each image, a vector of 512 values was obtained by passing the image through a series of Gabor filters across eight orientations and four spatial frequencies, and windowing the filtered images along an 8 x 8 grid. Each vector represents the image in terms of the spatial frequencies and orientations present at the different positions across the image. A cross-validation procedure was used to determine how similar individual images were to the average of each condition. GIST descriptors were averaged across all but one of the images within each condition. These average descriptors were then compared with each unique image creating within- and between-condition correlations for each combination of expression and viewpoint. The correlation values for the GIST descriptor across the different conditions were represented in a correlation matrix. This correlation matrix was also correlated with the average MVPA similarity matrix for each ROI and then used as a regressor for the fMRI data in the same way as the behavioural data from Experiments 1 and 2.

6.4 Results

6.4.1 Experiments 1

In Experiment 1 participants were asked to rate the perceptual similarity of expressions across pairs of images, whilst the factor they weren't rating (viewpoint) was changing across the two images. This experiment aimed to assess the impact of expression on the perception of the viewpoint. Average ratings of similarity for expression can be seen in Figure 6.2.



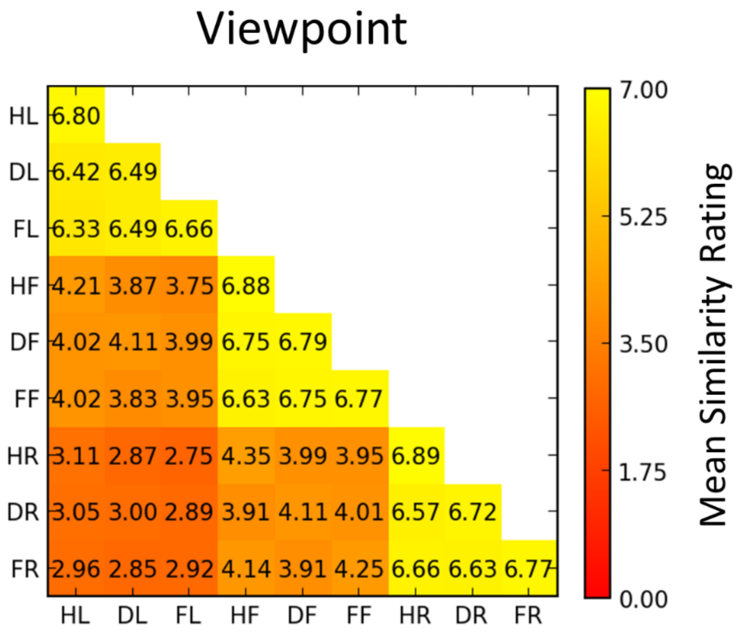
	HL	DL	FL	HF	DF	FF	HR	DR	FR
HL	6.84								
DL	1.53	6.38							
FL	1.58	2.77	5.64						
HF	6.52	1.51	1.35	6.72					
DF	1.53	6.03	2.39	1.65	6.35				
FF	1.38	2.26	5.07	1.44	2.71	5.58			
HR	6.66	1.75	1.61	6.64	1.60	1.48	6.74		
DR	1.44	6.33	2.79	1.38	6.32	2.35	1.70	6.42	
FR	1.57	2.99	5.47	1.52	2.52	5.51	1.49	2.82	5.58

Figure 6.2. Matrices showing the average perceptual similarity ratings from experiment 1 (Expression). HL: happy left ¾ profile, DL: disgust left ¾ profile, FL: fear left ¾ profile, HF: happy frontal view, DF: disgust frontal view, FF: fear frontal view, HR: happy right ¾ profile, DR: disgust right ¾ profile, FR: fear right ¾ profile.

Figure 6.2 demonstrates the higher similarity ratings for the perceptual similarity of expression are seen when the two images shared the same expression, independent of the viewpoint orientations changing.

6.4.2 Experiment 2

In Experiment 2 participants performed the same task as in Experiment 1, except they were rating the perceptual similarity of viewpoints across pairs of images, whilst the expression was changing across the two images. This experiment aimed to assess the impact of viewpoint on the perception of the expression. The average ratings of similarity for the viewpoint combinations can be seen in Figure 6.3.



	HL	DL	FL	HF	DF	FF	HR	DR	FR
HL	6.80								
DL	6.42	6.49							
FL	6.33	6.49	6.66						
HF	4.21	3.87	3.75	6.88					
DF	4.02	4.11	3.99	6.75	6.79				
FF	4.02	3.83	3.95	6.63	6.75	6.77			
HR	3.11	2.87	2.75	4.35	3.99	3.95	6.89		
DR	3.05	3.00	2.89	3.91	4.11	4.01	6.57	6.72	
FR	2.96	2.85	2.92	4.14	3.91	4.25	6.66	6.63	6.77

Figure 6.3. Matrices showing the average perceptual similarity ratings from experiment 2 (Viewpoint). HL: happy left ¾ profile, DL: disgust left ¾ profile, FL: fear left ¾ profile, HF: happy frontal view, DF: disgust frontal view, FF: fear frontal view, HR: happy right ¾ profile, DR: disgust right ¾ profile, FR: fear right ¾ profile.

For Experiment 2, the higher similarity ratings for the perceptual similarity of viewpoint directions across the two images were seen when the two images shared the same viewpoint direction and was independent of changes in expression. This demonstrates that in terms of the perceptual similarity of expression and viewpoints, these perceptions do not appear to be altered by changes in the other factor.

6.4.3 Experiment 3

The aim of Experiment 3 was to ask whether there are distinct neural representations in key regions of the face processing network, to specific combinations of expressions and viewpoints. MVPA was used to assess the similarity in the neural patterns of response to the nine conditions 1) happy; left ¾ profile (HL); 2) happy frontal view (HF), 3) happy; right ¾ profile (HR), 4) fear; left ¾ profile (FL); 5) fear frontal view (FF), 6) fear; right ¾ profile (FR), 7) disgust; left ¾ profile (DL); 8) disgust frontal view (DF), 9) disgust; right ¾ profile (DR). Figure 6.4 shows the matrices of the correlations between the patterns of response for the nine conditions across the ROIs. The matrices demonstrate higher correlations between the conditions which share a viewpoint direction (e.g. happy frontal view, disgust frontal view and fear frontal view) in the OFA, FFA and STS however there is no clear pattern in the results for the amygdala and IFG (Figure 6.4).

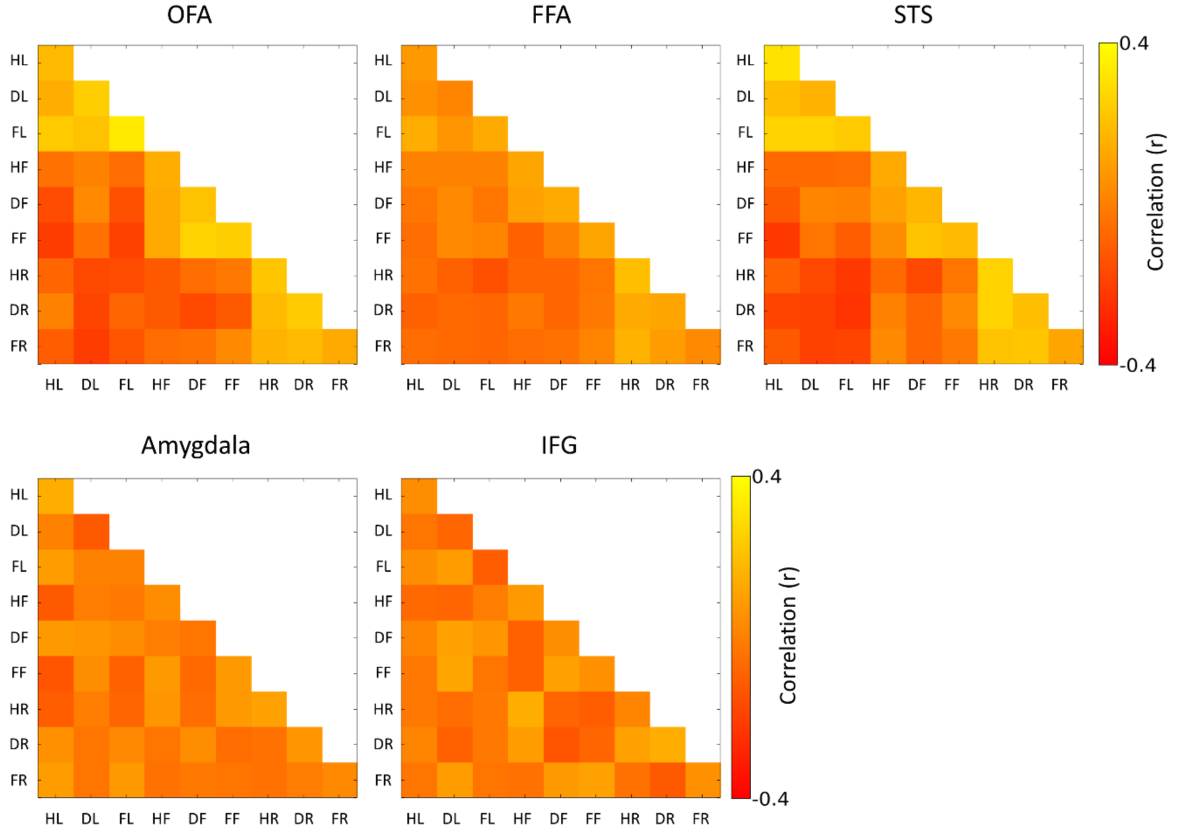


Figure 6.4. Correlation matrices demonstrating the similarity in the patterns of response between the nine conditions in each of the face-responsive ROIs. HL: happy left ¾ profile, DL: disgust left ¾ profile, FL: fear left ¾ profile, HF: happy frontal view, DF: disgust frontal view, FF: fear frontal view, HR: happy right ¾ profile, DR: disgust right ¾ profile, FR: fear right ¾ profile.

To determine the level of similarity between the behavioural similarity ratings and the neural data, correlations were made between the behavioural similarity matrices and the MVPA correlations matrix for each ROI. No significant correlations were seen between the patterns of neural response and the perceptual similarity ratings of expression from Experiment 1 (OFA: r = 0.07, p = .559, FFA: r = 0.12, p = .280, STS: 0.06, p = .595, IFG: r =

0.13, p = .249, Amygdala: r = 0.02, p = .887). However a strong positive correlation was seen between the patterns of neural response and the perceptual similarity ratings of viewpoint from Experiment 2 in the core regions (OFA: r = 0.89, p < .001, FFA: r = 0.78, p < .001, STS: r = 0.91, p < .001). Although no significant correlations were seen in the extended regions (IFG: r = 0.08, p = .475, Amygdala: r = 0.10, p = 0.346).

To determine if these effects were consistent across individuals, the behavioural data from Experiments 1 and 2 were used as a regressor for the fMRI MVPA matrices. These three models were entered into a simple linear regression to determine how much of the variance in the patterns of response could be explained by the perceptual similarity of expression and viewpoint. The results from the regression analysis can be seen in figure 6.5. The Viewpoint model explained a significant proportion of the variance in the core face-responsive regions (OFA: B = 0.45, p < .001, FFA: B = 0.22, p < .001, STS: 0.47, p < .001), but not the amygdala and IFG (IFG: B = 0.01, p = 603, Amygdala: B = 0.03, p = .146). The Expression model however, was not able to explain a significant proportion of the data in any of the ROIs (OFA: B = 0.02, p = .381, FFA: B = 0.03, p = .159, STS: B = 0.01, p = .546, IFG: B = 0.02, p = .397, Amygdala: B = 0.01, p = .809).

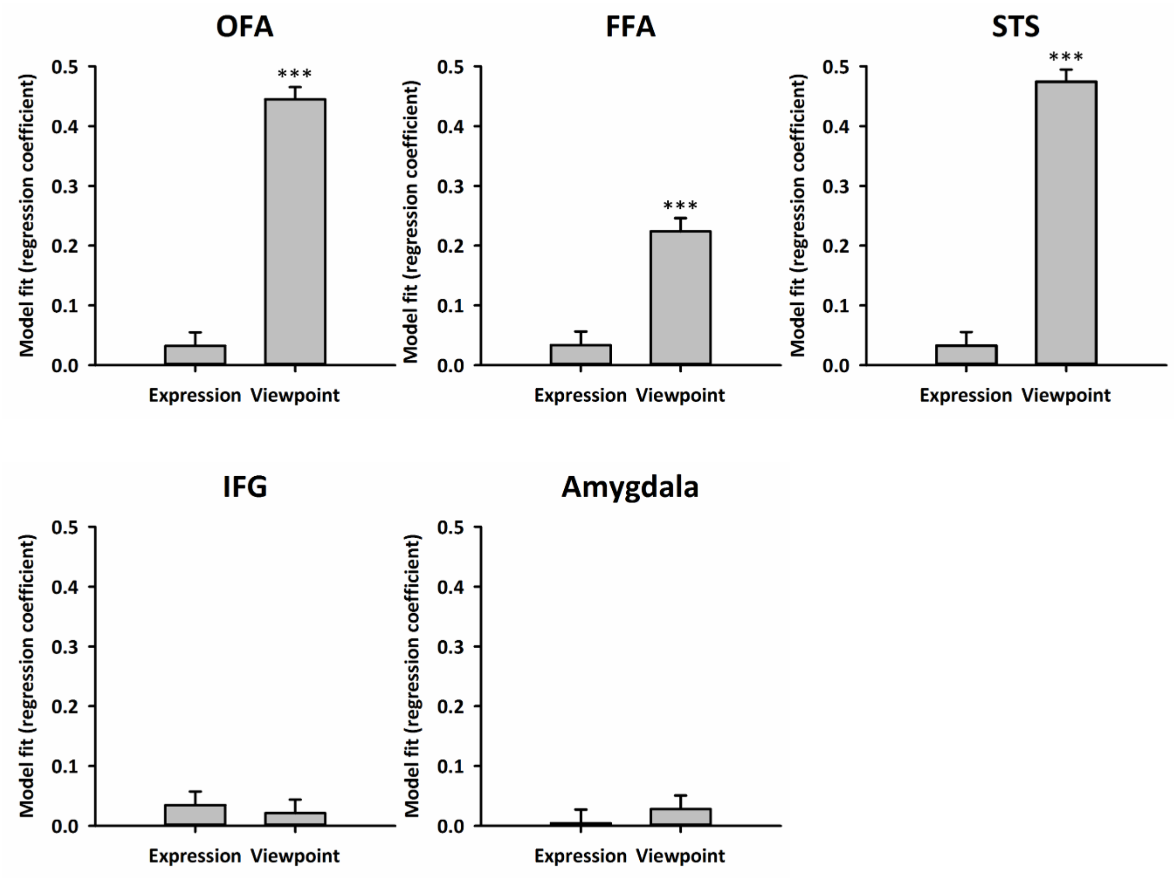


Figure 6.5. Results from the regression analysis demonstrating the amount of variance that can be explained by the Viewpoint and Expression models in the face-responsive regions.

This study then addressed whether the patterns of response from Experiment 3 could be explained by differences in the low-level image statistics for each of the viewpoint directions. The image statistics of each condition were analysed using a GIST descriptor (Oliva & Torralba, 2001). Figure 6.6 shows the correlations in image properties between conditions.

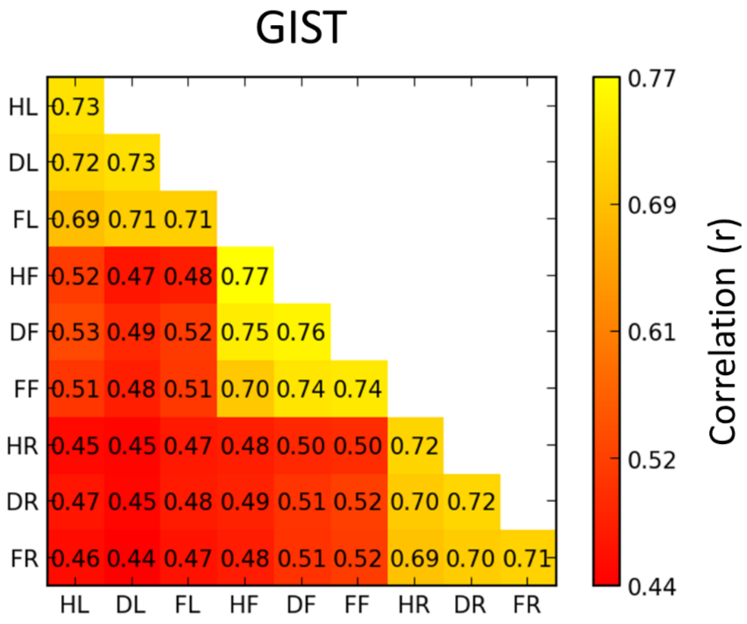


Figure 6.6. Correlation matrix demonstrating the similarity in image properties within- and between-conditions. It can clearly be seen that the images containing the same viewpoint direction have much more similar image statistics than images containing the same expression.

The average MVPA matrix for each ROI was then correlated with the GIST descriptor to demonstrate the level of correlation between the neural pattern of response, with the low-level image statics. A strong positive correlation was seen between the neural data and the low level image statistics in the core face-responsive regions (OFA: r = 0.88, p < .001, FFA: r = 0.75, p < .001, STS = 0.86, p < .001). However no significant correlation was seen in the extended regions (IFG: r = 0.06, p = .592, Amygdala: r = 0.10, p = .366).

To determine if these effects were consistent across individuals, the GIST descriptor matrix was also used as a regressor for the fMRI MVPA matrices, and entered as a model into a simple linear regression in the same manner as described above. The GIST descriptor was also able to explain a significant proportion of the variance in the core face-responsive regions (OFA: B = 0.44, p < .001, FFA: B = 0.22, p < .001, STS: 0.45, p < .001), but not in the extended regions of the IFG and the amygdala (IFG: B = 0.02, p = 408, Amygdala: B = 0.03, p = .229).

6.5 Discussion

The aim of this study was to explore the relative contribution of viewpoint and expression to patterns of response in the core face-responsive regions. There were distinct patterns of response to different viewpoints, but not to different expressions. This shows that the patterns of response in the core face-responsive regions are dominated by viewpoint. The high level of similarity between images that share the same viewpoint, suggests that the visual properties of the image are a dominant organising principle in these regions. This is supported by the positive correlation seen between the perceptual similarity ratings of viewpoint images, and a GIST descriptor of the images. Together these findings suggest that the viewpoint dominant responses may be linked to the underlying low-level properties of the images.

A GIST analysis of the stimuli used in this study demonstrated a high level of similarity between the images that was very similar to that seen in the neural patterns of response. A regression analysis used the results from Experiments 1 and 2, and the GIST analysis as models to see which could explain a significant proportion of the variance in the neural patterns of response. This demonstrated very similar performance for the GIST and behavioural data representing the perceptual similarity of viewpoint directions and strong positive correlations were seen between the GIST and viewpoint similarity ratings, with the patterns of response in the core regions. This suggests that the patterns of response to the combinations of expressions and viewpoints may be driven by the visual properties of the images.

Facial expressions produce much smaller visual changes than facial viewpoint directions and this may explain why both the GIST model and the perceptual similarity of viewpoint model were more strongly correlated, and performed much better in the regression analysis at explaining the variance in the neural patterns of response. A similar result was found in Chapter 5 where the visual properties of the images to various viewpoint directions, showed a great deal of similarity to the neural patterns of response to each of the viewpoint directions. Evidence from object perception has demonstrated that selectivity for object categories reflects underlying low-level image properties. This was also reflected in images of faces where a significant correlation between the neural patterns of response to faces was seen with the low-level properties of the images (Rice et al., 2014). Similar results have also been demonstrated with images of scenes (Watson, Hartley, & Andrews, 2014).

As discussed earlier, the pSTS has been identified as being involved in processing rigid movements, in addition to non-rigid movements of the face and is considered the core face-responsive region involved in the processing of changeable aspects of faces (Haxby et al., 2000). This would suggest the pSTS as being a candidate for a region that might have distinct neural representations for expression and viewpoint combinations. Unfortunately however, there was no evidence to suggest this and without further investigation it is difficult to explain why this might be, apart from the dominance of the larger change in image caused by viewpoint changes. However there is much research using univariate techniques that demonstrate clear responses to expressions (Allison et al., 2000; Baseler et al., 2013; Flack et al., 2015) and in Chapter 4 a distinct pattern of response was found to changes in expression. It is possible that the expression information is too fine-scale to be picked up by the MVPA technique, which is particularly useful for identifying course-scale representations. In addition the key difference between this study and that of Chapter 4 is that in the current study, a fixed expression was used throughout the block, whereas in Chapter 4 the expression changed across the block. This then may suggest that a univariate fMR adaptation technique might be better at helping us understand the representation of expression/viewpoint combinations.

The possibility of the expression information being too fine-scaled to be picked up by the technique used in this chapter, is somewhat supported by studies which found difficulty in decoding individual facial expressions (Srinivasan et al., 2016; Zhang et al., 2016). Although Srinivasan and colleagues were able to decode facial action units, they were unable to reliably decode individual facial expressions. The authors suggest this could be a result of the expression categories in the pSTS being coded at a finer scale, however they also consider that there may be greater variability across participants in how expressions are encoded.

In addition to the pSTS, the amygdala was another key candidate for a face-responsive region which demonstrates distinct coding of specific expression and viewpoint combinations. The amygdala in particular has been demonstrated to have a clear role in the processing of threat and univariate analyses has demonstrated the role of rigid changes of the face (e.g. eye gaze) in modulating the amygdala's response to facial expressions (Adams et al., 2003). Unfortunately this study did not produce any clear data for the amygdala and so it cannot be concluded from these results whether there is a distinct representation for these expression/viewpoint combinations. Given the results from the univariate data, it again suggests that this technique may not be appropriate for drawing out these kinds of responses. Indeed previous research has demonstrated differential responses to facial expressions within different sub-regions of the amygdala (Whalen et al., 2001) and so these types of responses could be lost in the use of a multivariate technique which for this study, simultaneously assess the response from large numbers of voxels.

Specific combinations of expressions and viewpoint have varying levels of relevance depending upon the social context. For example, a fearful expression facing away from you is likely to be of concern – you cannot see what the threat may be, however a fearful face looking directly at you is perhaps of less concern, as the individual may be fearful of you. In Experiments 1 and 2, participants had to rate the perceptual similarity of pairs of images in terms of their expression (Exp. 1) and viewpoint (Exp. 2). Participants showed no evidence of adjusting their perception of the expression or viewpoint, despite changes in the other factor. For example, the perceptual similarity of happy faces were not significantly affected by a change in viewpoint across the two images.

In conclusion, this study has demonstrated patterns of neural response to specific combinations of expressions and viewpoints, are dominated by the coding of the viewpoint direction. Results suggest this finding may reflect the underling low-level visual properties of the image, which is consistent with previous literature on the topographic representation of objects and scenes.

Chapter 7 – Summary and Conclusions

The information available from a human face is of great value. Not only can you tell key basic information such as what gender the person is and their approximate age, you can also determine whether they are familiar or unfamiliar to you. You can also determine information that is more changeable, such as how they are feeling and what they are paying attention to. These judgements can be made based on a person's facial expression and facial viewpoint and this can be done independently of recognition – you don't need to know who the person is, to make a reasonable judgment of where they're looking and how they're feeling. These judgements allow us to form relationships with other humans and enable effective interactions.

The human brain contains a network of regions that have been demonstrated to be face-responsive. These regions make up a core and extended network of regions purported to have differing roles in the processing of facial information. This thesis aimed to further our understanding of how facial expression and facial viewpoint are represented in these regions. Specifically whether these changeable aspects of faces have distinct representations, or whether they have overlapping representations in these face-responsive regions. The aims of this thesis were: 1) to investigate the neural representation of facial expression in the core and extended regions of the face processing network, 2) to investigate the neural representation of facial viewpoint in these same regions, and 3) to compare the representation of expression and viewpoint and determine whether these changes have distinct representations.

7.1 How are facial expressions and facial viewpoint represented in the brain?

Facial expressions are thought to be processed independently of invariant aspects of faces such as facial identity. This pathway is proposed (Haxby et al., 2000) to begin in the OFA and then proceed directly to the pSTS. From here, this information is further processed in regions in the extended system, such as the amygdala. The experiments in this thesis support the role of the pSTS in the processing of facial expression, however they also implicate a number of other regions.

The first experimental chapter (Chapter 3) used fMR adaptation to address how facial expression is represented in face-responsive areas of the human brain, with a specific focus on the pSTS. Behavioural experiments have shown that we process faces in a holistic way. That is, we represent the face as a perceptual whole rather than a collection of independent features. This phenomenon has been demonstrated using the face composite effect, which has supported holistic processing for both facial identity (Rossion 2013; Young, Hellawell, & Hay, 1987) and facial expression (Calder et al., 2000). Few studies have examined how the pSTS codes facial expression, and this chapter aimed to probe this coding by asking whether this region represents facial expression in a holistic or feature-based way.

The results from the pSTS show that this region is sensitive to any change in facial expression. The pSTS released from adaptation when either the top half or bottom half of the expression changed, and interestingly this did not require the face to be in a face-like configuration – the pSTS response was of roughly equal magnitude whether the top and bottom halves of the facial expression were aligned or misaligned. This relates well to a study that saw the pSTS respond in a linear fashion to changes in expression (Harris et al., 2012). Interestingly, the only region which processed facial expression in a holistic manner was the IFG. This region has been implicated in the processing of facial expressions by a number of studies, and is becoming more widely considered to be part of the face processing network (Carr, Iacoboni, Dubeau, Mazziotta, & Lenzi, 2003; Dapretto et al., 2006; Ishai, Schmidt, & Boesiger, 2005).

This thesis also aimed to address the proposed distinction between the processing of facial expressions and facial viewpoint. As described above, Haxby et al.'s (2000) model proposes separate neural pathways for the processing of these changeable aspects of faces and invariant aspects of faces, such as facial identity. This division in processing is also represented in behavioural models of face perception (Bruce & Young, 1986), where the analysis of changeable aspects of faces (expressions) and identity recognition are performed separately. This raises the question of how changeable aspects of faces are represented neurally. The results from Chapter 4, aimed to investigate this representation, and ask specifically whether there are distinct neural patterns for changes in expression and viewpoint and whether these are mediated by changes in facial identity. Experiment 1 demonstrated clear and distinct patterns of neural response to changes in facial expression in the OFA, FFA and pSTS. This suggests there are populations of voxels, which are selective for changes in facial expression, which are separate to those processing facial viewpoint. These neural patterns did not appear to be modulated by facial identity changes, with patterns of response to facial expression from the same identity, and from changing identities, being very similar. Experiment 2 took this further, by addressing the degree to which expression and viewpoint were represented in these regions. Experiment 2 was able to demonstrate that the proportions of voxels which respond to changes in expression and viewpoint vary across the face-responsive regions. The proportion of voxels in the regions that significantly respond to changes in expression increase from the OFA, to the FFA through to the pSTS, which is more classically considered to be involved in expression processing. However, in the IFG and amygdala, all voxels responded only to changes in expression, and not to changes in viewpoint. While this certainly supports previous literature implicating the amygdala in the processing of expression, the role of the IFG has been less clear. This experiment provides clear support for the role of the IFG in processing changes in facial expression.

Changes in an individual's viewpoint convey information regarding a person's focus of attention, and more basic information such as the presence of something salient in the environment. Humans are adept at discriminating between small changes in viewpoint direction (Wilson et al., 2000) and single cell studies have provided evidence of neurons which respond to specific viewpoint directions in macaques (Perrett et al., 1991) and indirect evidence of this in humans Fang and He (2005).

Chapter 5 aimed to ask whether there are distinct neural patterns of response in face-responsive regions, for individual viewpoint directions. Viewpoint-specific representations were found in the core regions of the OFA, FFA and pSTS. This relates well to work by Axelrod and Yovel (2012) who were able to decode facial viewpoint in the OFA, FFA and STS. However, it was clear from the results in Chapter 5, that there was more to the viewpoint representations than distinct coding for each direction. There was also similarity between the patterns of neural response to symmetrical viewpoint directions (e.g. left and right profile).

An interesting progression was found in the way viewpoint was represented across the core face-responsive regions. The ability to distinguish between different viewpoint directions (left vs right) was higher in the OFA, where a model based on ability to code viewpoint direction was able to explain the data well, and to a similar level as a GIST descriptor of the low-level image properties. However, in the FFA and pSTS, similar patterns of response to symmetrical viewpoint directions became more apparent. This confusion between symmetrical viewpoints could potentially constitute an efficient processing shortcut. Previous research has suggested that similarity in neural response to symmetric viewpoint directions may represent partial viewpoint invariance. This partial viewpoint invariance could form an intermediate processing stage from viewpoint specific coding, to full viewpoint invariance (Axelrod & Yovel, 2012; Kietzmann et al., 2012). When considering where we see symmetry in the world, it is often in two halves of the same object. Therefore, there isn't necessarily a strong case for being able to distinguish between the two, and the similarity in neural responses perhaps captures this.

The final experimental chapter (Chapter 6) explored whether it was possible to decode specific combinations of facial expression and viewpoints (e.g. happy left, disgust front, fear right). There is mixed evidence to suggest that facial expressions can be decoded in various face-responsive regions. Whilst some studies find it is possible to decode facial expressions in regions such as the STS (Said et al., 2010), other studies have found this more challenging (Srinivasan et al., 2016; Zhang et al., 2016). Srinivasan et al. (2016) were able to demonstrate distinct patterns of response to facial action units, which reflect the underlying facial musculature system and are used for describing and distinguishing between different facial movements, e.g. facial expressions. However they weren't able to reliably decode facial expression categories. Zhang et al. (2016) were able to distinguish between emotional expressions versus neutral, but not individual expressions within the emotional category (fearful, angry, happy). There is stronger evidence to show that facial viewpoints can be decoded in higher level visual areas (Axelrod & Yovel, 2012; Kietzmann et al., 2012), and this was also demonstrated in Chapter 5. So, whilst there is reasonable evidence that both facial expressions and facial viewpoints can be decoded in various higher level visual areas, the ability to decode specific combinations of the two, had not been addressed.

The final experimental chapter (Chapter 6) aimed to determine whether there are distinct representations of expression and viewpoint combinations. In addition, this chapter aimed to determine the relative dominance of expression and viewpoint in the neural representation of the core and extended regions. Distinct patterns of response to different expressions were not found. However there was a high level of similarity in images which shared the same viewpoint. This shows that the patterns of response were largely dominated by viewpoint. A strong positive correlation was seen between the neural patterns of response with the underlying low-level visual properties present in the images. These results suggest that the representation of specific expression and viewpoint combinations may be based on the underlying statistical properties of the images, rather than higher level categories such as expression and viewpoint.

Overall, this thesis has made a significant contribution to our understanding of face perception. The demonstration of holistic processing of facial expressions helps to build a clearer picture of the IFG’s involvement in face perception and gives an interesting and novel resolution to the neural basis of holistic expression processing. Demonstrating distinct patterns of response to facial expressions and facial viewpoint across the core regions has added to the building body of literature suggesting we may need to think differently about how the brain represents faces and other objects. The research presented in this thesis gives weight to the theory that categories of objects are represented as distinct and overlapping patterns of response, rather than as peak responses in specific regions. This significantly aids our understanding of how the brain could represent such immeasurable numbers of objects in a limited area of cortex. The regions studied in the previous chapters were part of a popular and well tested neurological model of face perception (Haxby et al., 2000). Whilst the previous chapters very much support the roles of the core regions of the OFA, FFA and STS in face processing, they suggest that the distinct roles assigned to these regions may not be so clear cut, as expression and viewpoint information appears to be represented to some level in all of these regions.

From a wider perspective, this research demonstrates a significant degree of similarity in the way individuals process facial information. The consistency in the patterns of response across individuals suggest that to some extent, we process facial expressions and facial viewpoints in the same way. This makes a lot of intuitive sense, as we know our brains have roughly the same developmental trajectories. The fact that we can sum over tens of thousands of voxels across multiple individuals and find common representations of information, is incredible. This could potentially lead to a common map of how the human brain represents various objects and concepts, and therefore a map to demonstrate normal functioning for comparison with individuals with brain lesions.

7.2 Reflections and future directions

The experiments in this thesis have produced a fascinating insight into the workings of the core and extended face processing regions. There were studies that produced very interesting results, particularly those of Chapters 3 and 5, however there were other results that were somewhat disappointing, for example, the lack of expression decoding in Chapter 6. Therefore, it has likely been a very typical scientific journey! The results of Chapter 6 led to a lot of consideration about the visual processing (as opposed to more higher level processing) of facial expression. From a visual perspective, changes in viewpoint produce a much larger visual change than facial expressions. However, the brain needs to be able to compensate for this, and in relation to Chapter 5, this information is also important to ignore in some cases. On the other hand, changes in expression are small visual changes, but constitute very significant changes for social communication. In essence, big changes in the image are not necessarily psychologically significant, and small changes in the image are not necessarily psychologically insignificant. The scale of the visual change, and the scale of its importance are not necessarily the same. The psychological relevance and interpretation of these changes are likely to be processed further downstream in the processing pathway, perhaps beginning in the extended regions and beyond.

The symmetric viewpoint responses seen in Chapter 5, open up many possibilities for future research directions. This chapter adds to the growing evidence suggesting we might have a process by which we go from a viewpoint specific representation of a face, to partial viewpoint invariance, through to full viewpoint invariance. Chapter 5 uses unfamiliar faces, and so if this theory were to be the case, full viewpoint invariance is unlikely to be achieved in a short fMRI experiment with unfamiliar faces. Some degree of partial viewpoint invariance is feasible however, due to seeing multiple repetitions of the same identity. A good test of how this theory relates to human behaviour, would be to use a face recognition memory paradigm. In real life we learn faces from multiple encounters and we become familiar with a person over time. Face recognition memory studies use repeated exposure to multiple images (Longmore & Young, 2008). This format would allow a good test of whether learning an individual would show an advantage for symmetrical viewpoint directions.

In terms of furthering this research using fMRI, an interesting complementary study would be to use an adaptation paradigm. If the patterns of response to symmetric viewpoint directions are considered the same, or very similar, then it may be possible to see adaptation to a specific viewpoint, without seeing a release from adaptation to the symmetrical opposite of this viewpoint.

7.3 Overall summary

The aim of this thesis was to further our understanding of the neural representation of facial expressions and facial viewpoint. This thesis has made a significant contribution to the scientific literature by presenting evidence to support the existence of distinct patterns of response to expression and viewpoint in core face-responsive regions and that these changeable aspects of faces are represented by an overall neural pattern across these regions. In addition this thesis was also able to demonstrate evidence of viewpoint symmetry representations and importantly, this thesis has also given weight to the evidence suggesting the IFG has a significant role in the processing of facial expressions.

References

Adams, R. B., Gordon, H. L., Baird, A. A., Ambady, N., & Kleck, R. E. (2003). Effects of Gaze on Amygdala Sensitivity to Anger and Fear Faces. *Science*, *300*(5625), 1536–1536.

Adams, R. B., & Kleck, R. E. (2003). Perceived Gaze direction and the processing of facial displays of emotion. *Psychological Science*, *14*(6), 644–647.

Adolphs, R., Gosselin, F., Buchanan, T., Tranel, D., Schyns, P., & Damasion, A. (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature*, *433*(7021), 68–72.

Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. R. (1995). Fear and the human amygdala. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *15*(9), 5879–91.

Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. (1994). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature*, *372*.

Adolphs, R., Tranel, D., Hamann, S., Young, A. W., Calder, A. J., Phelps, E. A., Anderson, Al, Lee, G. P. & Damasio, A. R. (1999). Recognition of facial emotion in nine individuals with bilateral amygdala damage. *Neuropsychologia*, *37*(10), 1111–1117.

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends in cognitive sciences*, *4*(7), 267–278.

Andrews, T. J., Davies-Thompson, J., Kingstone, A., & Young, A. W. (2010). Internal and external features of the face are represented holistically in face-selective regions of visual cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *30*(9), 3544–52.

Andrews, T. J., & Ewbank, M. P. (2004). Distinct representations for facial identity and changeable aspects of faces in the human temporal lobe. *NeuroImage*, *23*(3), 905–13.

Asghar, A. U. R., Chiu, Y.-C., Hallam, G., Liu, S., Mole, H., Wright, H., & Young, A. W. (2008).

An amygdala response to fearful faces with covered eyes. *Neuropsychologia*, *46*(9), 2364–70.

Axelrod, V., & Yovel, G. (2012). Hierarchical processing of face viewpoint in human visual cortex. *Journal of Neuroscience*, *32*(7), 2442–2452.

Axelrod, V., & Yovel, G. (2013). The challenge of localizing the anterior temporal face area: A possible solution. *NeuroImage*, *81*, 371-380.

Baseler, H. A., Harris, R. J., Young, A. W., & Andrews, T. J. (2014). Neural Responses to Expression and Gaze in the Posterior Superior Temporal Sulcus Interact with Facial Identity. *Cerebral Cortex*, *24*(3), 737.

Baudouin, J.-Y., & Humphreys, G. W. (2006). Compensatory strategies in processing facial emotions: evidence from prosopagnosia. *Neuropsychologia*, *44*(8), 1361–9.

Bird, G., & Cook, R. (2013). Mixed emotions: the contribution of alexithymia to the emotional symptoms of autism. *Translational Psychiatry*, *3*(7), e285.

Booth, M. C. A., & Rolls, E. T. (1998). View-invariant representations of familiar objects by neurons in the inferior temporal visual cortex. *Cerebral Cortex*, *8*(6), 510–523.

Bouvier, S. E. (2005). Behavioral Deficits and Cortical Damage Loci in Cerebral Achromatopsia. *Cerebral Cortex*, *16*(2), 183–191.

Broks, P., Young, A. W., Maratos, E. J., Coffey, P. J., Calder, A. J., Isaac, C. L., … Hadley, D. (1998). Face processing impairments after encephalitis: Amygdala damage and recognition of fear. *Neuropsychologia*, *36*(1), 59–70.

Bruce, V. (1986). Influences of familiarity on the processing of faces. *Perception*, *15*(4), 387–397.

Bruce, V., & Young, A. W. (1998). *In the eye of the beholder: the science of face perception.* Oxford: Oxford University Press.

Bruce, V., & Young, A. W. (2012). *Face Perception*.

Burton, A. M., Jenkins, R., & Schweinberger, S. R. (2011). Mental representations of familiar faces. *British journal of psychology (London, England : 1953)*, *102*(4), 943–58.

Calder, A. J., Lawrence, A. D., & Young, A. W. (2001). Neuropsychology of fear and loathing. *Nature reviews. Neuroscience*, *2*(5), 352–63.

Calder, A. J. (1996). Facial Emotion Recognition after Bilateral Amygdala Damage. *Cognitive Neuropsychology*, *3*(5), 699–745.

Calder, A. J., Beaver, J. D., Winston, J. S., Dolan, R. J., Jenkins, R., Eger, E., & Henson, R. N. a. (2007). Separate coding of different gaze directions in the superior temporal sulcus and inferior parietal lobule. *Current biology : CB*, *17*(1), 20–5.

Calder, A. J., & Jansen, J. (2005). Configural coding of facial expressions: The impact of inversion and photographic negative. *Visual Cognition*, *12*(3), 495–518.

Calder, A. J., & Nummenmaa, L. (2007). Face cells: separate processing of expression and gaze in the amygdala. *Current biology : CB*, *17*(10), R371–2.

Calder, A. J., & Young, A. W. (2005). Understanding the recognition of facial identity and facial expression. *Nature reviews. Neuroscience*, *6*(8), 641–51.

Calder, A. J., Young, A. W., Keane, J., & Dean, M. (2000). Configural information in facial expression perception. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(2), 527–551.

Calvert, G. A. (2001). Crossmodal processing in the human brain: insights from functional neuroimaging studies. *Cerebral cortex (New York, N.Y. : 1991)*, *11*(12), 1110–23.

Campbell, R., Heywood, C. a, Cowey, a, Regard, M., & Landis, T. (1990). Sensitivity to eye gaze in prosopagnosic patients and monkeys with superior temporal sulcus ablation. *Neuropsychologia*, *28*(11), 1123–1142.

Carlin, J. D., & Calder, A. J. (2012). The neural basis of eye gaze processing. *Current opinion in neurobiology*, 10–15.

Carlin, J. D., Calder, A. J., Kriegeskorte, N., Nili, H., & Rowe, J. B. (2011). A Head View-Invariant Representation of Gaze Direction in Anterior Superior Temporal Sulcus. *Current Biology*, *21*(21), 1817–1821.

Carlin, J. D., Rowe, J. B., Kriegeskorte, N., Thompson, R., & Calder, A. J. (2012). Direction-sensitive codes for observed head turns in human superior temporal sulcus. *Cerebral cortex*, *22*(4), 735–44.

Carr, L., Iacoboni, M., Dubeau, M.-C., Mazziotta, J. C., & Lenzi, G. L. (2003). Neural mechanisms of empathy in humans: a relay from neural systems for imitation to limbic areas. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(9), 5497–502.

Cohen Kadosh, K., Henson, R. N. a, Cohen Kadosh, R., Johnson, M. H., & Dick, F. (2010). Task-dependent activation of face-sensitive cortex: an fMRI adaptation study. *Journal of cognitive neuroscience*, *22*(5), 903–917.

Cook, R., Brewer, R., Shah, P., & Bird, G. (2013). Alexithymia, Not Autism, Predicts Poor Recognition of Emotional Facial Expressions. *Psychological Science*, *24*(5), 723–732.

Dapretto, M., Davies, M. S., Pfeifer, J. H., Scott, A. A., Sigman, M., Bookheimer, S. Y., & Iacoboni, M. (2006). Understanding emotions in others: mirror neuron dysfunction in children with autism spectrum disorders. *Nature Neuroscience*, *9*(1), 28–30.

Davies-Thompson, J., & Andrews, T. J. (2012). Intra- and interhemispheric connectivity between face-selective regions in the human brain, 3087–3095.

Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*.

Ekman, P., Friesen, W. V, O'Sullivan, M., Chan, a, Diacoyanni-Tarlatzis, I., Heider, K., … Ricci-Bitti, P. E. (1987). Universals and cultural differences in the judgments of facial expressions of emotion. *Journal of personality and social psychology*, *53*(4), 712–7.

Ekman, P., & Friesen, W. (1986). A New Pan-Cultural Facial Expression of Emotion 1. *Motivation and Emotion*, *10*(2), 159–168.

Ekman, P., & Heider, K. G. (1988). The universality of a contempt expression: A replication. *Motivation and Emotion*, *12*(3), 303–308.

Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin*, *128*(2), 203–235.

Engell, A. D., & Haxby, J. V. (2007). Facial expression and gaze-direction in human superior temporal sulcus. *Neuropsychologia*, *45*(14), 3234–41.

Ewbank, M. P., & Andrews, T. J. (2008). Differential sensitivity for viewpoint between familiar and unfamiliar faces in human visual cortex. *NeuroImage*, *40*(4), 1857–70.

Fairhall, S. L., & Ishai, A. (2007). Effective connectivity within the distributed cortical network for face perception. *Cerebral Cortex*, *17*(10), 2400–2406.

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

Fang, F., Murray, S. O., & He, S. (2007). Duration-dependent fMRI adaptation and distributed viewer-centered face representation in human visual cortex. *Cerebral Cortex*, *17*(6), 1402–1411.

Feinstein, J. S., Adolphs, R., Damasio, A., & Tranel, D. (2011). The human amygdala and the induction and experience of fear. *Current biology : CB*, *21*(1), 34–8.

Feinstein, J. S., Buzza, C., Hurlemann, R., Follmer, R. L., Dahdaleh, N. S., Coryell, W. H., … Wemmie, J. A. (2013). Fear and panic in humans with bilateral amygdala damage. *Nature Neuroscience*, *16*(3), 270–272.

Flack, T. R., Andrews, T. J., Hymers, M., Al-Mosaiwi, M., Marsden, S. P., Strachan, J. W. A., Trakulpipat, C., Wang, L., Wu, T. & Young, A. W. (2015). Responses in the right posterior superior temporal sulcus show a feature-based response to facial expression. *Cortex*, *69*, 14–23.

Fox, C. J., Moon, S. Y., Iaria, G., & Barton, J. J. S. (2009). The correlates of subjective perception of identity and expression in the face network: An fMRI adaptation study. *NeuroImage*, *44*(2), 569–580.

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

Friesen, C. K., & Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin & Review*, *5*(3), 490–495.

Ganel, T., Valyear, K. F., Goshen-Gottstein, Y., & Goodale, M. A. (2005). The involvement of the “fusiform face area” in processing facial expression. *Neuropsychologia*, *43*(11), 1645–1654.

Gauthier, I., & Tarr, M. J. (1997). Becoming a “Greeble” expert: Exploring mechanisms for face recognition. *Vision Research*, *37*(12), 1673–1682.

Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (1999). Activation of the middle fusiform “face area” increases with expertise in recognizing novel objects. *Nature neuroscience*, *2*(6), 568–73.

Gilbert, C., & Bakan, P. (1973). Visual asymmetry in perception of faces. *Neuropsychologia*, *11*(3), 355–362.

Gobbini, M. I., & Haxby, J. V. (2007). Neural systems for recognition of familiar faces. *Neuropsychologia*, *45*(1), 32–41.

Grill-Spector, K., Knouf, N., & Kanwisher, N. (2004). The fusiform face area subserves face perception, not generic within-category identification. *Nature Neuroscience*, *7*(5), 555–562.

Guo, K., & Shaw, H. (2015). Face in profile view reduces perceived facial expression intensity: An eye-tracking study. *Acta psychologica*, *155*, 19–28.

Hadjikhani, N., Hoge, R., Snyder, J., & de Gelder, B. (2008). Pointing with the eyes: the role of gaze in communicating danger. *Brain and cognition*, *68*(1), 1–8.

Hagan, C. C., Woods, W., Johnson, S., Calder, A. J., Green, G. G. R., & Young, A. W. (2009). MEG demonstrates a supra-additive response to facial and vocal emotion in the right superior temporal sulcus. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(47), 20010–5.

Hanke, M., Halchenko, Y. O., Sederberg, P. B., Hanson, S. J., Haxby, J. V, & Pollmann, S. (2009). PyMVPA: a Python Toolbox for Multivariate Pattern Analysis of fMRI Data. *Neuroinformatics*, *7*(1), 37–53.

Harris, R. J., Young, A. W., & Andrews, T. J. (2012). Morphing between expressions dissociates continuous from categorical representations of facial expression in the human brain. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(51), 21164–9.

Harris, R. J., Young, A. W., & Andrews, T. J. (2014). Dynamic stimuli demonstrate a categorical representation of facial expression in the amygdala. *Neuropsychologia*, *56*(1), 47–52.

Haxby, J. V., Hoffman, E., & Gobbini, M. (2000). The distributed human neural system for face perception. *Trends in cognitive sciences*, *4*(6), 223–233.

Haxby, J. V. (2012). Multivariate pattern analysis of fMRI: the early beginnings. *NeuroImage*, *62*(2), 852–5.

Haxby, J. V, Gobbini, M. I., Furey, M. L., Ishai, a, Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science (New York, N.Y.)*, *293*(5539), 2425–30.

Haynes, J.-D., & Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature reviews. Neuroscience*, *7*(7), 523–34.

Hess, U., Adams, R. B., & Kleck, R. E. (2007). Looking at You or Looking Elsewhere: The Influence of Head Orientation on the Signal Value of Emotional Facial Expressions. *Motivation and Emotion*, *31*(2), 137–144.

Hoffman, E. A., & Haxby, J. V. (2000). Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nature neuroscience*, *3*(1), 80–4.

Humphreys, G. W., Donnelly, N., & Riddoch, M. J. (1993). Expression is computed separately from facial identity, and it is computed separately for moving and static faces: neuropsychological evidence. *Neuropsychologia*, *31*(2), 173–181.

Ishai, A. (2002). Visual Imagery of Famous Faces: Effects of Memory and Attention Revealed by fMRI. *NeuroImage*, *17*(4), 1729–1741.

Ishai, A. (2008). Let’s face it: It's a cortical network. *NeuroImage*, *40*(2), 415–419.

Ishai, A., Schmidt, C. F., & Boesiger, P. (2005). Face perception is mediated by a distributed cortical network. *Brain Research Bulletin*, *67*(1-2), 87–93.

Ishai, A., Ungerleider, L. G., & Haxby, J. V. (2000). Distributed neural systems for the generation of visual images. *Neuron*, *28*(3), 979–90.

Izard, C. E., & Haynes, O. M. (1988). On the form and universality of the contempt expression: A challenge to Ekman and Friesen’s claim of discovery. *Motivation and Emotion*, *12*(1), 1–16.

Jenkinson, M., & Smith, S. (2001). A global optimisation method for robust affine registration of brain images, *5*, 143–156.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *17*(11), 4302–11.

Kanwisher, N., & Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *361*(1476), 2109–2128.

Kietzmann, T. C., Swisher, J. D., Konig, P., & Tong, F. (2012). Prevalence of Selectivity for Mirror-Symmetric Views of Faces in the Ventral and Dorsal Visual Pathways. *Journal of Neuroscience*, *32*(34), 11763–11772.

Kietzmann, T., Gert, A., & König, P. (2015). Representational dynamics of facial viewpoint encoding: Head orientation, viewpoint symmetry, and front-on views. *Journal of Vision*, *15*(12), 750.

Kleinke, C. L. (1986). Gaze and eye contact: a research review. *Psychological bulletin*, *100*(1), 78–100.

Kourtzi, Z., Erb, M., Grodd, W., & Bülthoff, H. H. (2003). Representation of the perceived 3-D object shape in the human lateral occipital complex. *Cerebral Cortex*, *13*(9), 911–920.

Kriegeskorte, N., Mur, M., & Bandettini, P. (2008). Representational similarity analysis - connecting the branches of systems neuroscience. *Frontiers in systems neuroscience*, *2*(November), 4.

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition & Emotion*, *24*(8), 1377–1388.

Lee, L. C., Andrews, T. J., Johnson, S. J., Woods, W., Gouws, A., Green, G. G. R., & Young, A. W. (2010). Neural responses to rigidly moving faces displaying shifts in social attention investigated with fMRI and MEG. *Neuropsychologia*, *48*(2), 477–90.

Leveroni, C. L., Seidenberg, M., Mayer, A. R., Mead, L. a, Binder, J. R., & Rao, S. M. (2000). Neural Systems Underlying the Recognition of Familiar and Newly Learned Faces. *The Journal of Neuroscience*, *20*(2), 878–886.

Liu, J., Harris, A., & Kanwisher, N. (2009). Perception of face parts and face configurations: an FMRI study. *Journal of cognitive neuroscience*, *22*(1), 203–211.

Longmore, C. A., & Young, A. W. (2008). Learning Faces From Photographs, *34*(1), 77–100.

Mattavelli, G., Sormaz, M., Flack, T., Asghar, A. U. R., Fan, S., Frey, J., … Andrews, T. J. (2014). Neural responses to facial expressions support the role of the amygdala in processing threat. *Social Cognitive and Affective Neuroscience*, *9*(11), 1684–1689.

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*, *6*(6), 255–260.

McKone, E., & Kanwisher, N. (2005). Does the human brain process objects of expertise like faces? *From Monkey Brain to Human Brain: A Fyssen Foundation Symposium*.

Morris, J. S., deBonis, M., & Dolan, R. J. (2002). Human Amygdala Responses to Fearful Eyes. *NeuroImage*, *17*(1), 214–222.

Morris, J. S., Frith, C. D., Perrett, D. I., Rowland, D., Young, A. W., Calder, A. J., & Dolan, R. J. (1996). A differential neural response in the human amygdala to fearful and happy facial expressions. *Nature*.

Narumoto, J., Okada, T., Sadato, N., Fukui, K., & Yonekura, Y. (2001). Attention to emotion modulates fMRI activity in human right superior temporal sulcus. *Brain research. Cognitive brain research*, *12*(2), 225–31.

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

Nichols, D. F., Betts, L. R., & Wilson, H. R. (2010). Decoding of faces and face components in face-sensitive human visual cortex. *Frontiers in Psychology*, *1*(July), 1–13.

Nummenmaa, L., & Calder, A. J. (2009). Neural mechanisms of social attention. *Trends in cognitive sciences*, *13*(3), 135–43.

Oliva, A., & Torralba, A. (2001). Modeling the shape of the scene: A holistic representation of the spatial envelope. *International Journal of Computer Vision*, *42*(3), 145–175.

Or, C. C. F., & Wilson, H. R. (2010). Face recognition: Are viewpoint and identity processed after face detection? *Vision Research*, *50*(16), 1581–1589.

Palermo, R., Willis, M. L., Rivolta, D., McKone, E., Wilson, C. E., & Calder, A. J. (2011). Impaired holistic coding of facial expression and facial identity in congenital prosopagnosia. *Neuropsychologia*, *49*(5), 1226–35.

Peirce, J. W. (2007). PsychoPy--Psychophysics software in Python. *Journal of neuroscience methods*, *162*(1-2), 8–13.

Pelphrey, K., Viola, R., & McCarthy, G. (2004). When Strangers Pass Processing of Mutual and Averted Social Gaze in the Superior Temporal Sulcus. *Psychological Science*, *15*(9), 598–604.

Perrett, D. I., Oram, M. W., Harries, M. H., Bevan, R., Hietanen, J. K., Benson, P. J., & Thomas, S. (1991). Viewer-centred and object-centred coding of heads in the macaque temporal cortex. *Experimental Brain Research*, *86*(1), 159–173.

Perrett, D. I., Smith, P. A., Potter, D. D., Mistlin, A. J., Head, A. S., Milner, A. D., & Jeeves, M. A. (1985). Visual cells in the temporal cortex sensitive to face view and gaze direction. *Proceedings of the Royal Society of London. Series B, Containing papers of a Biological character. Royal Society (Great Britain)*, *223*(1232), 293–317.

Pitcher, D., Charles, L., Devlin, J. T., Walsh, V., & Duchaine, B. (2009). Triple dissociation of faces, bodies, and objects in extrastriate cortex. *Current biology : CB*, *19*(4), 319–24.

Pitcher, D., Garrido, L., Walsh, V., & Duchaine, B. C. (2008). Transcranial magnetic stimulation disrupts the perception and embodiment of facial expressions. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *28*(36), 8929–33.

Pitcher, D., Walsh, V., & Duchaine, B. (2011). The role of the occipital face area in the cortical face perception network. *Experimental Brain Research*, *209*(4), 481–493.

Pitcher, D., Walsh, V., Yovel, G., & Duchaine, B. (2007). TMS evidence for the involvement of the right occipital face area in early face processing. *Current biology : CB*, *17*(18), 1568–73.

Prazak, E. R., & Burgund, E. D. (2014). Keeping it real: Recognizing expressions in real compared to schematic faces. *Visual Cognition*, *22*(5), 737–750.

Psalta, L., Young, a. W., Thompson, P., & Andrews, T. J. (2014). The Thatcher Illusion Reveals Orientation Dependence in Brain Regions Involved in Processing Facial Expressions. *Psychological Science*, *25*(1), 128–136.

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

Ramirez, F. M., Cichy, R. M., Allefeld, C., & Haynes, J.-D. (2014). The Neural Code for Face Orientation in the Human Fusiform Face Area. *Journal of Neuroscience*, *34*(36), 12155–12167.

Rice, G. E., Watson, D. M., Hartley, T., & Andrews, T. J. (2014). Low-Level Image Properties of Visual Objects Predict Patterns of Neural Response across Category-Selective Regions of the Ventral Visual Pathway. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *34*(26), 8837–44.

Rossion, B. (2013). The composite face illusion: A whole window into our understanding of holistic face perception. *Visual Cognition*, *21*(2), 139–253.

Rotshtein, P., Henson, R. N. A., Treves, A., Driver, J., & Dolan, R. J. (2005). Morphing Marilyn into Maggie dissociates physical and identity face representations in the brain. *Nature Neuroscience*, *8*(1), 107–113.

Said, C. P., Moore, C. D., Engell, A. D., & Haxby, J. V. (2010). Distributed representations of dynamic facial expressions in the superior temporal sulcus. *Journal of Vision*, *10*(2010), 1–12.

Schiltz, C., Dricot, L., Goebel, R., & Rossion, B. (2010). Holistic perception of individual faces in the right middle fusiform gyrus as evidenced by the composite face illusion. *Journal of Vision*, *10*(2), 25.1–16.

Schiltz, C., & Rossion, B. (2006). Faces are represented holistically in the human occipito-temporal cortex. *NeuroImage*, *32*(3), 1385–94.

Srinivasan, R., Golomb, J. D., & Martinez, A. M. (2016). A Neural Basis of Facial Action Recognition in Humans. *Journal of Neuroscience*, *36*(16), 4434–4442.

Steeves, J. K. E., Culham, J. C., Duchaine, B. C., Pratesi, C. C., Valyear, K. F., Schindler, I., … Goodale, M. a. (2006). The fusiform face area is not sufficient for face recognition: evidence from a patient with dense prosopagnosia and no occipital face area. *Neuropsychologia*, *44*(4), 594–609.

Tarr, M. J., & Gauthier, I. (2000). FFA: a flexible fusiform area for subordinate-level visual processing automatized by expertise. *Nature neuroscience*, *3*(8), 764–9.

Tong, F., & Pratte, M. S. (2012). Decoding patterns of human brain activity. *Annual review of psychology*, *63*, 483–509.

Watson, D. M., Hartley, T., & Andrews, T. J. (2014). Patterns of response to visual scenes are linked to the low-level properties of the image. *NeuroImage*, *99*, 402–410.

Whalen, P. J. (2004). Human Amygdala Responsivity to Masked Fearful Eye Whites. *Science*, *306*(5704), 2061–2061.

Whalen, P. J., Shin, L. M., McInerney, S. C., Fischer, H., Wright, C. I., & Rauch, S. L. (2001). A functional MRI study of human amygdala responses to facial expressions of fear versus anger. *Emotion*, *1*(1), 70–83.

Wicker, B., Michel, F., Henaff, M. a, & Decety, J. (1998). Brain regions involved in the perception of gaze: a PET study. *NeuroImage*, *8*(2), 221–227.

Wilson, H. R., Wilkinson, F., Lin, L. M., & Castillo, M. (2000). Perception of head orientation. *Vision Research*, *40*(5), 459–472.

Winston, J. S., Henson, R. N. a, Fine-Goulden, M. R., & Dolan, R. J. (2004). fMRI-adaptation reveals dissociable neural representations of identity and expression in face perception. *Journal of neurophysiology*, *92*(3), 1830–9.

Wright, T. M., Pelphrey, K. a, Allison, T., McKeown, M. J., & McCarthy, G. (2003). Polysensory interactions along lateral temporal regions evoked by audiovisual speech. *Cerebral cortex (New York, N.Y. : 1991)*, *13*(10), 1034–43.

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception*, *16*(6), 747–59.

Young, A. W., McWeeny, K. H., Hay, D. C., & Ellis, A. W. (1986). Matching familiar and unfamiliar faces on identity and expression. *Psychological Research*, *48*(2), 63–68.

Young, A. W., Newcombe, F., Haan, E. H. F. De, Small, M., & Hay, D. C. (1993). Face perception after brain injury. *Brain*, *116*, 941–959.

Young A. W., Perrett D.I., Calder AJ, S. R., & P, E. (2002). *Facial Expressions of Emotion: Stimuli and Tests (FEEST) (Thames Valley Test Company, Bury St Edmunds, England). Psychology*.

Zhang, H., Japee, S., Nolan, R., Chu, C., Liu, N., & Ungerleider, L. G. (2016). Face-selective regions differ in their ability to classify facial expressions. *NeuroImage*, *130*, 77–90.