The neural coding of properties shared by faces, bodies and objects

Dissertation

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

der Mathematisch-Naturwissenschaftlichen Fakultät

und

der Medizinischen Fakultät

der Eberhard-Karls-Universität Tübingen

vorgelegt

von

Celia Foster

aus Manchester, England

Januar 2020

Tag der mündlichen Prüfung: 7th September 2020

Dekan der Math.-Nat. Fakultät: Prof. Dr. J. Fortágh

Dekan der Medizinischen Fakultät: Prof. Dr. B. Pichler

1. Berichterstatter: Prof. Andreas Bartels

2. Berichterstatter: PD Dr. Marc Himmelbach

Prüfungskommission: Prof. Andreas Bartels

PD Dr. Marc Himmelbach

Prof. Michael Black

Prof. Hong Yu Wong

Erklärung / Declaration:

Ich erkläre, dass ich die zur Promotion eingereichte Arbeit mit dem Titel:
„The neural coding of properties shared by faces, bodies and objects“
selbständig verfasst, nur die angegebenen Quellen und Hilfsmittel benutzt und wörtlich oder inhaltlich übernommene Stellen als solche gekennzeichnet habe. Ich versichere an Eides statt, dass diese Angaben wahr sind und dass ich nichts verschwiegen habe. Mir ist bekannt, dass die falsche Abgabe einer Versicherung an Eides statt mit Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft wird.

I hereby declare that I have produced the work entitled “The neural coding of properties shared by faces, bodies and objects”, submitted for the award of a doctorate, on my own (without external help), have used only the sources and aids indicated and have marked passages included from other works, whether verbatim or in content, as such. I swear upon oath that these statements are true and that I have not concealed anything. I am aware that making a false declaration under oath is punishable by a term of imprisonment of up to three years or by a fine.

Tübingen, den ....

Abstract

Previous studies have identified relatively separated regions of the brain that respond strongly when participants view images of either faces, bodies or objects. The aim of this thesis was to investigate how and where in the brain shared properties of faces, bodies and objects are processed. We selected three properties that are shared by faces and bodies, shared categories (sex and weight), shared identity and shared orientation (i.e. facing direction). We also investigated one property shared by faces and objects, the tendency to process a face or object as a whole rather than by its parts, which is known as holistic processing. We hypothesized that these shared properties might be encoded separately for faces, bodies and objects in the previously defined domain-specific regions, or alternatively that they might be encoded in an overlapping or shared code in those or other regions. In all of studies in this thesis, we used fMRI to record the brain activity of participants viewing images of faces and bodies or objects that showed differences in the shared properties of interest. We then investigated the neural responses these stimuli elicited in a variety of specifically localized brain regions responsive to faces, bodies or objects, as well as across the whole-brain. Our results showed evidence for a mix of overlapping coding, shared coding and domain-specific coding, depending on the particular property and the level of abstraction of its neural coding. We found we could decode face and body categories, identities and orientations from both face- and body-responsive regions showing that these properties are encoded in overlapping brain regions. We also found that non-domain specific brain regions are involved in holistic face processing. We identified shared coding of orientation and weight in the occipital cortex and shared coding of identity in the early visual cortex, right inferior occipital cortex, right parahippocampal cortex and right superior parietal cortex, demonstrating that a variety of brain regions combine face and body information into a common code. In contrast to these findings, we found evidence that high-level visual transformations may be predominantly processed in domain-specific regions, as we could most consistently decode body categories across image-size and body identity across viewpoint from body-responsive regions. In conclusion, this thesis furthers our understanding of the neural coding of face, body and object properties and gives new insights into the functional organisation of occipitotemporal cortex.

Contents

1. Synopsis

One of the most impressive abilities of our visual system is that we can not only easily detect whether a scene contains a person, as well as other objects, but we can also easily identify specific properties of this person. When we see various images of people we can recognise their identity, determine whether they belong to particular categories (e.g. whether they are male or female) and ascertain whether their body, or direction of gaze, is oriented toward us or not. Our ability to recognize these properties is particularly impressive when we consider that there can be a huge variety in the low-level visual information present in different images. For example, we can identify the same person in different images despite variations of the illumination and viewpoint, or when the body is partially occluded. Despite all of these sources of interference, our visual system can discriminate even quite subtle properties, for example when we distinguish the identities of two individuals that have very similar facial features.

How does the visual system achieve these impressive abilities? Visual scientists have investigated the neural circuits thought to underlie visual processing of people and have also attempted to replicate these abilities in computational models. Deep convolutional neural networks, computational models inspired by the organisation of the nervous system containing multiple layers of nodes that are trained to perform recognition and categorization of images using large training datasets, have been shown to be able to achieve impressive and human-like object recognition abilities (Krizhevsky, Sutskever, & Hinton, 2012; LeCun, Bengio, & Hinton, 2015). However, these models are still limited when compared to the human visual system. They lack its flexibility, require large training datasets, and small perturbations, which are hardly visible to human observers, have been shown to cause large misclassifications of objects by these networks (Szegedy et al., 2014). More importantly, we don't know exactly what computational circuits within these models give rise to their recognition abilities. In order to develop better computational models, and to understand what goes wrong in humans who have impairments in recognition abilities (Behrmann & Avidan, 2005), it is important for us to better understand how visual information is processed and represented in the brain, and how these processes allow our visual system to robustly recognize features of people (and more generally objects) under a variety of visual conditions.

1.1. Category-responsive brain regions

Light entering the eye first activates the photoreceptors of the retina, and following this the neural circuits of the retina begin to process this visual information (Masland, 2012). This processed visual information is then sent to the lateral geniculate nucleus, and then further on to the primary visual cortex, known as V1. Visual information is further processed in V1, where the neurons have been found to respond to specific orientations of bars of light (Hubel & Wiesel, 1962). Beyond V1, it is thought that further processing of visual information in occipital and temporal cortex underlies our visual abilities to detect and process properties of people and objects. Early work investigating neural responses in the macaque inferotemporal cortex identified neurons with complex object-related response properties, for example, neurons responsive to the shape of a hand (Gross, Roch-Miranda, & Bender, 1972).

In 1990, the discovery that magnetic resonance imaging could be used to detect blood oxygenation level changes related to neural activity (Ogawa, Lee, Nayak, & Glynn, 1990; Ogawa, Lee, Kay, & Tank, 1990) led to a new method for neuroscientists to investigate regional changes in human brain activity (Logothetis, 2002). This method, known as functional magnetic resonance imaging (fMRI), made it more feasible for neuroscientists to investigate the functional organisation of the healthy human brain, and made it easier to test how multiple brain regions respond to many different kinds of stimuli. One influential early study using this method identified a region in the fusiform gyrus that responded stronger when participants viewed images of faces than when they viewed many other kinds of visual stimuli, including objects, scrambled face images, houses and hands (Kanwisher, McDermott, & Chun, 1997). This region is commonly referred to as the fusiform face area (FFA). By showing that this region responded stronger to faces than to many other different kinds of visual stimuli, the authors argued that this provided strong evidence that this region is specifically involved in processing faces. Following this finding, subsequent fMRI studies identified more regions, in addition to the FFA, that responded strongly to faces compared to other kinds of stimuli. These regions include the occipital face area (OFA) located in the lateral occipital cortex (Gauthier, Tarr, et al., 2000), the superior temporal sulcus (STS) (Halgren et al., 1999; Haxby et al., 1999) and the anterior temporal face area (ATFA) (Rajimehr, Young, & Tootell, 2009; Tsao, Moeller, & Freiwald, 2008), as well as many

other regions that are hypothesised to be part of a distributed face-responsive network (Haxby, Hoffman, & Gobbini, 2000; Ishai, 2008).

Beyond faces, fMRI studies have identified brain regions showing strong responses to other kinds of specific categories. A region of the lateral and ventral occipital cortex, known as the lateral occipital complex (LOC), has been shown to respond higher when participants view images of objects compared to when they view a variety of texture or visual noise images (Grill-Spector, Kushnir, Edelman, Itzchak, & Malach, 1998; Grill-Spector, Kushnir, Hendler, et al., 1998; Malach et al., 1995). Brain regions showing higher responses when participants view images of scenes, compared to when they view images of objects or faces, have been identified in the parahippocampal cortex (known as the parahippocampal place area, PPA) (Epstein & Kanwisher, 1998), the retrosplenial cortex (RSC) (Epstein, 2008) and the transverse occipital cortex (TOS, also referred to as occipital place area, OPA) (Kamps, Julian, Kubilius, Kanwisher, & Dilks, 2016). A region in the left fusiform gyrus has been shown to respond when participants visually process words, known as the visual word form area (VWFA) (Cohen et al., 2000; McCandliss, Cohen, & Dehaene, 2003). Two brain regions have been identified that are more responsive to headless bodies (where all face information is removed) compared to objects or scenes, known as the extrastriate body area (EBA) (Downing, Jiang, Shuman, & Kanwisher, 2001) and the fusiform body area (FBA) (Peelen & Downing, 2005). Altogether these studies show that there are many regions across a large area of occipitotemporal cortex that respond preferentially to specific visual categories. Interestingly, these category-responsive regions show remarkable consistency in their locations across different participants (Op de Beeck, Pillet, & Ritchie, 2019).

1.2. Factors contributing to the overall arrangement of category-responsive regions

Several studies have tried to determine whether there is an overarching explanation for why we have this particular arrangement of category-responsive brain regions. Several factors have been proposed to contribute to the overall spatial organization and selectivity of these cortical regions. One key factor that has been proposed is animacy. Strong differences between the neural representations of animate and inanimate stimuli have been found in the occipitotemporal cortex of both humans and monkeys (Kriegeskorte et al., 2008). In relation to the previously defined category-responsive regions, this distinction is demonstrated in the separation between regions encoding scenes and objects, and those

encoding faces and bodies. For example, the scene-responsive PPA and ventral portion of the object-responsive LOC (known as pFs) are located medially to the face- and body-responsive FFA and FBA on the fusiform gyrus. Interestingly, although there are many differences in low-level visual features between animate and inanimate stimuli, for example differences in spatial frequency, there is evidence that these low-level differences cannot fully explain the neural responses to animate and inanimate stimuli. Distinctions in the neural responses to animate versus inanimate stimuli have been demonstrated in stimuli controlled for low-level visual differences (Bracci & Op de Beeck, 2016). Furthermore, category-specific neural responses in occipitotemporal cortex have been shown to be elicited by highly abstract representations of these categories. For example, the FFA has been shown to respond to line drawings of faces and two-tone 'Mooney faces' (Kanwisher, Tong, & Nakayama, 1998; Loffler, Yourganov, Wilkinson, & Wilson, 2005; Tong, Nakayama, Moscovitch, Weinrib, & Kanwisher, 2000) and similarly the body-responsive EBA and FBA have been shown to respond to stick-figure drawings of bodies (Downing et al., 2001; Peelen & Downing, 2005). Thus, it seems there is a strong distinction between the neural coding of animate and inanimate stimuli that cannot be explained solely by co-occurring differences in low-level visual properties.

Another factor that has been proposed to contribute to the overall arrangement of category-responsive brain regions is the location in the visual field where these different categories are predominately processed. Early visual areas have been shown to map visual information in a retinotopic organisation (Wandell, Dumoulin, & Brewer, 2007). As different kinds of object categories tend to take up smaller (e.g. faces) or larger (e.g. buildings) areas of the visual field, this means that different neurons in these early visual regions process these different object categories. Furthermore, as we tend to fixate certain objects, such as faces (Yarbus, 1967), it has been proposed that small fixated objects would be processed more by neurons responding to the centre of the visual field, whereas buildings and scenes that cover large portions of the visual field would be more likely to be processed by neurons encoding the peripheral visual field (Levy, Hasson, Avidan, Hendler, & Malach, 2001). Neuroimaging studies have found evidence for such central-peripheral biases in category-responsive regions of the occipitotemporal cortex, for example the FFA was found to respond more to central than peripheral stimuli, whereas the PPA was found to respond

more to peripheral then central stimuli (Hasson, Harel, Levy, & Malach, 2003; Hasson, Levy, Behrmann, Hendler, & Malach, 2002; Levy et al., 2001). Other work has demonstrated an additional contribution of the real-world size of objects to the neural representations in occipitotemporal cortex (Konkle & Oliva, 2012). In this study, the authors found that inanimate objects with similar real-world size were represented in nearby clusters of occipitotemporal cortex, regardless of the image-sizes of these objects shown during the experiment. Furthermore, the authors found that regions responsive to large objects were also activated when participants imagined small objects at a large scale (or vice versa), demonstrating that this finding seemed to be related to the participants' perception of the size of the object. In combination, these studies demonstrate that the functional organisation of occipitotemporal cortex is influenced by objects sizes and their corresponding retinal locations in our natural visual environment.

Studies have also demonstrated that the neural organisation of occipitotemporal cortex is not solely driven by visual features. The most striking example of this is the finding that a similar arrangement of category-specific responses is evoked by other sensory modalities. For example, tactile recognition of faces and objects has been shown to evoke activity in occipitotemporal areas that are also activated during visual recognition of faces and objects, and furthermore these regions are also activated during tactile recognition of faces and objects in congenitally blind individuals, ruling out that these results could be due to visual imagery of the tactile stimuli (Pietrini et al., 2004). Subsequent studies have found further evidence for similarities between tactile and visual category-responses in occipitotemporal cortex. For both sighted and blind individuals, higher activity has been shown in the EBA during haptic identification of hands compared to objects (Kitada et al., 2014), and activity in the middle temporal gyrus has been shown during haptic recognition of facial expressions (Kitada et al., 2013). Furthermore, the VWFA has been shown to be activated when congenitally blind participants read via touch using Braille (Reich, Szwed, Cohen, & Amedi, 2011). A similar overlap in activation has also been found between visual category-responsive regions and auditory category responses in both sighted and blind participants. Listening to category-specific words elicits similar distinctions between animate and inanimate stimuli (Mahon, Anzellotti, Schwarzbach, Zampini, & Caramazza, 2009), and PPA has been found to be more active when participants listen to large

nonmanipulable object words as compared to small object or animal words (He et al., 2013). Furthermore, natural sounds related to faces, bodies, scenes and objects were found to elicit a similar arrangement of responses as compared to visual stimuli in ventral temporal cortex, for both sighted and blind participants (van den Hurk, Van Baelen, & Op de Beeck, 2017). In sum, these multisensory studies demonstrate that the arrangement of category-responsive regions in occipitotemporal cortex is, to some extent, influenced by multisensory processing or an innate map, rather than being solely driven by visual experience during development.

1.3. Functional processing within category-responsive regions

How does neural activity in the category-responsive brain regions relate to the detection and processing of specific properties of these categories? It has been proposed that the arrangement of category-responsive brain regions could be somehow necessary or optimal for our visual system to solve complex problems, such as recognising specific people and objects (Grill-Spector & Weiner, 2014). Furthermore, it has been proposed that the reason for the existence of multiple, separated regions responding to the same category (e.g. the face-responsive regions in occipital, fusiform, superior temporal and anterior temporal cortex) could be because these regions encode representations of different properties of that category (Haxby et al., 2000).

Early fMRI studies investigating face-responses found evidence for such a functional separation between different face-responsive brain regions. In an experiment where participants attended to either the identity or gaze direction of faces images, higher activity was found in the STS when participants attended to gaze as compared to identity, whereas the opposite was true for the fusiform gyrus, where responses were higher when participants attended to identity as compared to eye gaze (Hoffman & Haxby, 2000). The authors proposed a neural separation between changeable (e.g. facial expression, gaze direction) and unchangeable (e.g. identity) aspects of faces (Haxby et al., 2000). This proposal corresponded with a model of face perception based on psychological findings that also proposed a separation of different aspects of face processing (Bruce & Young, 1986). Several subsequent neuroimaging studies found further evidence for this neural distinction of face processing. Neural responses to emotion expression (R. J. Harris, Young, & Andrews, 2012; Srinivasan, Golomb, & Martinez, 2016; Zhang et al., 2016), gaze direction (Carlin &

Calder, 2013; Carlin, Calder, Kriegeskorte, Nili, & Rowe, 2011) and facial motion (Schultz & Pilz, 2009) have been found in the STS, whereas neural responses to face identity have been found in the FFA and anterior temporal face area (Grill-spector, Knouf, & Kanwisher, 2004; Nasr & Tootell, 2012; Rotshtein, Henson, Treves, Driver, & Dolan, 2005). These studies show there is considerable evidence for functional differences between the face-responsive regions in the STS and fusiform gyrus.

Functional differences related to part-based versus configural processing have also been proposed between more posterior and more anterior regions with the same category-responsive preference. The OFA has been found to respond more to face parts than to whole faces (Arcurio, Gold, & James, 2012) and to contain a map of face-features (Henriksson, Mur, & Kriegeskorte, 2015). Furthermore, one study found that the FFA responds to the correct configuration of face parts, but found no evidence for a sensitivity to face-part configuration in the OFA (Liu, Harris, & Kanwisher, 2010). For bodies, the EBA has been shown to be more sensitive to body parts than the FBA (Taylor, Wiggett, & Downing, 2007), and a topographic body-part map has been identified that corresponds with the location of the EBA (Orlov, Makin, & Zohary, 2010). For scenes, the PPA and RSC have been shown to respond to the correct configuration of scene components (e.g. correct positioning of walls relative to the floor), whereas the TOS does not show this distinction, despite showing responses to the scene components (Kamps et al., 2016). Altogether, these studies suggest that more posterior category-responsive regions, that are closer to the early visual cortex, encode local elements of the categories, whereas more anterior category-responsive regions respond to correct overall configurations of the categories.

Several studies have also identified functional differences related to viewpoint-dependence versus viewpoint-invariance between more posterior and more anterior category-responsive regions. For faces, an electrophysiological recording study in macaque monkeys identified posterior to anterior differences in the viewpoint-dependence of face responses in the macaque face-responsive patches (Freiwald & Tsao, 2010). The authors found that responses in the two most posterior face-responsive patches (lateral and middle fundus, ML and MF) were viewpoint-dependent, whereas moving anteriorly responses in the next patch (anterior lateral, AL) showed mirror-symmetric face responses (where neurons responded equally to mirror-symmetric views) and finally responses in the most

anterior patch (anterior medial, AM) showed viewpoint-invariant face responses. Thus, this study showed a posterior to anterior hierarchy in face-responses from viewpoint-dependent to viewpoint-invariant, with mid-level regions showing a partial invariance. Human fMRI studies have also found evidence for a similar hierarchy in the human face-responsive regions. Viewpoint-invariant face identity responses have been identified in the ATFA and FFA (Anzellotti, Fairhall, & Caramazza, 2014; Guntupalli, Wheeler, & Gobbini, 2017), and some studies have also found evidence for mirror-symmetric face responses in the FFA and STS (Axelrod & Yovel, 2012; Flack, Harris, Young, & Andrews, 2019; Guntupalli et al., 2017; Kietzmann, Swisher, König, & Tong, 2012), however this evidence is debated due to methodological concerns (Ramírez, 2018; Ramírez, Cichy, Allefeld, & Haynes, 2014). Viewpoint-specific face responses have been found in the OFA, FFA and early visual areas (Axelrod & Yovel, 2012; Flack et al., 2019; Guntupalli et al., 2017; Ramírez et al., 2014). There is some evidence of a similar distinction between viewpoint-dependent and viewpoint-independent responses for bodies and objects. For bodies, electrophysiological recordings in macaque monkeys found that body-responses in a more posterior body-responsive patch were more viewpoint-dependent than body responses in a more anterior body-responsive patch (Kumar, Popivanov, & Vogels, 2019). For objects and scenes, similar neural responses to mirror-reversed objects images were found in be the anterior portion of the LOC (pFs) but not in the posterior portion (LO), and similar responses to mirror-reversed scene images were found in the PPA, but not in the TOS or RSC (Dilks, Julian, Kubilius, Spelke, & Kanwisher, 2011). These results suggest that neurons in the more anterior object and scene regions show a tolerance to the image changes induced by mirror-reversal, whereas neurons in the more posterior object and scene regions respond to specific viewpoints of objects and scenes. In combination, these studies suggest an overall increase in viewpoint-invariant responses from posterior to more anterior category-responsive regions, which has been proposed to be a general processing principle that allows for viewpoint-invariant recognition of people and objects (Freiwald & Tsao, 2010).

1.4. Are the functional responses in category-responsive regions domain-specific?

Many studies assume that neural processing of properties of a particular category will occur in the defined regions responsive to this particular category (e.g. coding of face properties in face-responsive regions). However, several studies have challenged the

hypothesis that category-responsive regions are strictly selective for their preferred category. One influential study found that patterns of activity in regions outside of the strongly category-responsive regions could be used to identify that category (Haxby et al., 2001). The authors proposed that representations of faces and objects are distributed in overlapping patterns across ventral temporal cortex. Another study found that category-responsive regions contain patterns of neural activity that could distinguish between subordinate groups of a non-preferred category (Op de Beeck, Brants, Baeck, & Wagemans, 2010). This study provided further evidence that neural responses in category-responsive regions are not entirely category-selective.

The category-specificity of the FFA for faces has also been challenged. The FFA has also been shown to respond to objects of expertise, for example when bird experts view images or birds, or car experts view images of cars (Gauthier, Skudlarski, Gore, & Anderson, 2000; Yaoda Xu, 2005). Furthermore, neural activity in the FFA evoked by novel objects has been show to increase when participants are trained to develop expertise with these novel objects (Gauthier, Tarr, Anderson, Skudlarski, & Gore, 1999). Therefore, the authors proposed that rather than being category-selective for faces, the FFA may be involved in subordinate-level processing of any class of objects that a person has expertise with.

These studies provide strong evidence that neural responses in category-responsive regions are not fully domain-specific. Importantly, this suggests that the neural coding of functional properties of these categories may be more widespread than was originally thought. Furthermore, these findings highlight a problem with the current methodical approaches being used to investigate the neural coding of properties associated with these categories. In order to boost statistical power, many studies specifically localize category-responsive regions and then specifically investigate the neural coding of properties associated with this category only within these category-responsive regions (e.g. investigating face properties in face-responsive regions) and not within a wider variety of brain regions. In such studies, regions outside of these category-responsive regions that also respond to the property would not be detected. It is possible that this approach could lead to biases and missing information about the functional properties of occipitotemporal cortex, as responses that are not domain-specific would not be detected.

In this thesis, we aimed to explore the neural coding of face, body and object properties both within and outside of the domain-specific regions corresponding to each category. This approach allowed us to investigate whether the neural coding of a variety of properties was encoded in domain-specific category-responsive regions, or encoded in more distributed brain regions. We specifically chose to investigate properties that are shared by faces and bodies or faces and objects. The rationale was, that if category-responsive regions process only properties of their preferred categories, then we would find neural representations of these properties in domain-specific regions (e.g. properties of faces in face-responsive regions). In contrast, if the neural coding of functional properties is more general, then a shared property should be encoded in an overlapping or shared neural representation.

1.5. The neural coding of properties shared by faces and bodies or objects

The overall aim of this thesis was to investigate the neural coding of shared properties of faces and bodies or objects. In Chapters 2, 3 and 4, we investigated the neural coding of three shared properties of faces and bodies. Faces and bodies are particularly suitable for investigating shared properties. As a whole person contains a face and a body, there are many properties specific to a particular person that are visible from both their face and body. Despite the number of shared properties, neuroimaging studies have identified prominent, yet separated brain regions in the occipital and fusiform cortex that respond when participants view a face or a body (Peelen & Downing, 2005; Premereur, Taubert, Janssen, Vogels, & Vanduffel, 2016; Schwarzlose, Baker, & Kanwisher, 2005). Thus, faces and bodies contain many shared properties that can be investigated and well-defined brain regions, which can be easily localized in almost all participants, where we can investigate the neural coding of these shared properties.

In Chapters 5 and 6, we investigated the neural processes involved with holistic processing of faces and objects. We chose to compare holistic processing of faces and objects, rather than faces and bodies, as behavioural studies investigating whether bodies are processed holistically have found mixed results (Bauser, Suchan, & Daum, 2011; Bonemei, Costantino, Battistel, & Rivolta, 2018; Brandman & Yovel, 2010; A. Harris, Vyas, & Reed, 2016; Robbins & Coltheart, 2012a; Willems, Vrancken, Germeys, & Verfaillie, 2014; Yovel, Pelc, & Lubetzky, 2010). Thus, it may be that bodies are only weakly processed

holistically, or are only processed holistically under certain circumstances. In contrast to these mixed findings, behavioural studies have found strong evidence of holistic processing of the objects we tested in Chapter 6, moreover evidence of holistic processing of these objects was found to be as strong as for faces (Zhao, Bülthoff, & Bülthoff, 2016). Thus, we considered that the neural processes involved with holistic processing of these objects would be more easily detectable than those involved with holistic processing of bodies and additionally would be more suitable to compare to the neural processes involved with holistic processing of faces.

We investigated the neural coding of four different shared properties in the present work: shared subcategories, shared identity, shared orientation and holistic processing. As it was not possible to investigate all possible shared properties of faces, bodies and objects, we selected the subset of shared properties to investigate in this thesis based on three criteria. The first criterion was how prominent the property was for faces and bodies or objects. We predicted that prominent properties would likely give strong neural responses that would be possible for us to detect using fMRI. We determined whether a property was prominent or not based on evidence from behavioural studies investigating the detection or processing of the particular property. The second criterion was the novelty in researching the property. As some properties have been more extensively studied than others, we aimed to select properties that have received less focus in order to maximise the contribution of this work to the overall understanding of face, body and object processing. The third criterion aimed to maximise the variety between the four properties we selected. Even though we could only investigate a subset of all the possible shared properties of faces, bodies and objects, we aimed to encompass some of the wide variety of kinds of shared properties in our selection. In combination, we believed that the properties selected with these criteria would show strong neural responses and allow for new insights into the neural processes underlying a variety of shared properties of faces, bodies and objects.

In Chapter 2, we investigated the neural coding of two subcategories shared by faces and bodies, sex and weight. We created images of faces and bodies that varied in sex (male or female) and weight (higher or lower weight), and validated that these images were perceived to belong to these subcategories via a perceptual rating behavioural experiment. We then conducted an fMRI study, where we recorded the brain activity of participants

viewing these images. Using multivoxel pattern analyses (MVPA), we investigated which brain regions contain separable patterns of activity evoked by male and female, and higher and lower weight stimuli. In a first set of analyses, we tested whether face subcategories would be encoded only in face-responsive brain regions, and body subcategories in body-responsive brain regions, or if there would be overlap in the regions encoding the face and body subcategories. We found we could decode body subcategories from both body- and face-responsive brain regions, but image-size invariant decoding was more specific to the body-responsive regions. We could decode the weight of faces from the FBA, again suggesting some overlap in the regions encoding shared subcategories. In a second set of analyses, we tested if any regions contain subcategory representations that could generalise across neural activity evoked by faces and bodies. We could decode weight across neural activity evoked by faces and bodies in the EBA, suggesting some abstract coding of weight in this region. In combination, this study shows that there is an overlap in the neural coding of face and body subcategories, as we could decode body subcategories from face-responsive regions and face subcategories form body-responsive regions. However, we also found evidence that image-size invariant coding may be more domain-specific. In combination, these results support a mix of shared and separated neural coding of subcategories, depending on the level of abstraction of the neural representation.

In Chapter 3, we investigated the neural coding of identities shared by the face and body. Although, psychological research has shown that we use information from both the face and body to identify people (Hahn, O'Toole, & Phillips, 2015; O'Toole et al., 2011; Rice, Phillips, Natu, An, & O'Toole, 2013; Rice, Phillips, & O'Toole, 2013; Robbins & Coltheart, 2012b), there has been little research into how the brain integrates identity information from the face and body. In this study, we trained participants to recognize three individuals from images of their face and body, and then recorded the participants' brain activity using fMRI as they viewed these images. We then used MVPA to investigate which brain regions contain separable patterns of neural activity evoked by the three identities. In a first set of analyses, we tested whether face identity and body identity are encoded separately in face- and body-responsive regions respectively or not. We found we could decode face identity from the face-responsive ATFA and the body-responsive EBA, and we could decode body identity from the body-responsive FBA and face-responsive OFA. These results suggest there

is a mixed coding of identity information in face- and body-responsive brain regions for both faces and bodies. In further analyses, we found we could decode body identity across different body viewpoints from the FBA but not consistently from any face-responsive regions. This suggests viewpoint-invariant body identity representations may be body-region specific. In a second set of analyses, we tested if any brain regions contain identity representations that could generalize across neural activity evoked by faces and bodies. We found we could decode identity in this abstract manner from the early visual cortex, right inferior occipital cortex, right parahippocampal cortex and right superior parietal cortex, but not from any of our face- or body-responsive ROIs. Altogether, these results show that information about both face and body identity is contained in both face- and body-responsive regions, but higher-level viewpoint-invariant identity representations may be domain-specific. Furthermore, abstract identity representations that generalise across face and body viewing were found to be encoded in several distributed brain regions outside of the commonly defined face- or body-responsive regions.

In Chapter 4, we investigated the neural coding of orientation directions shared by faces and bodies. We see faces and bodies from many different orientations (e.g. facing towards us) and the same orientation directions can apply to both faces and bodies. Most previous studies have separately investigated the neural coding of face and body orientation, therefore it is unknown whether face and body orientations are encoded in a shared or separated manner in occipitotemporal cortex. In this study, we recorded the brain activity of participants viewing images of faces and bodies from three orientations with fMRI. We then used MVPA to investigate which brain regions contained separable patterns of activity evoked by the three orientations. We found that the face-responsive OFA and body responsive EBA contain neural coding of both face and body orientation, and furthermore contained abstract neural coding of orientation that could generalize across neural activity evoked by faces and bodies. Furthermore, we found neural responses to face orientation, but not body orientation, in the FFA and FBA, suggesting that these fusiform-gyrus regions encode face-specific orientation information. In sum, these results show an abstract encoding of person orientation in the early face- and body-responsive regions of the occipital cortex. Furthermore, they also show a later face-specific coding of orientation

in the fusiform-gyrus, which interestingly was found in both face- and body-responsive fusiform regions.

In Chapters 5 & 6, we investigated which brain regions are involved in holistic processing of faces and non-expertise objects. Holistic processing refers to the tendency to perceive something as an indecomposable whole rather than by its component parts. Many behavioural studies have demonstrated that faces are processed holistically (Maurer, Le Grand, & Mondloch, 2002; Richler & Gauthier, 2014) and a recent study found strong evidence that certain non-expertise objects are also processed holistically (Zhao et al., 2016). Neuroimaging studies have linked holistic processing of faces and expertise objects to neural activity in the FFA (Gauthier & Tarr, 2002; Goffaux, Schiltz, Mur, & Goebel, 2013; Liu et al., 2010; Ross et al., 2018; Schiltz, Dricot, Goebel, & Rossion, 2010; Schiltz & Rossion, 2006; Wong, Palmeri, Rogers, Gore, & Gauthier, 2009), however recent behavioural work has found evidence that there may be additional processes involved in holistic processing of faces (Curby & Moerel, 2019). So far no studies have investigated which brain regions are involved in holistic processing of non-expertise objects. In these studies, we used fMRI to record the brain activity of participants performing composite-tasks with faces and non-expertise objects. We localized a variety of brain regions that we hypothesized could be involved in holistic processing, and then investigated the neural responses in these regions during the composite-task. In Chapter 5, we found evidence of neural responses related to holistic processing of faces in both the face-responsive FFA2 (an anterior component of the FFA) as well as in the object-responsive LOC and scene-responsive RSC, PPA and TOS. Interestingly we found that the neural responses related to holistic processing in the RSC and PPA correlated with participants behavioural responses in the composite-task, suggesting a direct relationship between neural activity in these regions and participant's behaviour. In Chapter 6, we did not find any significant neural responses related to holistic processing of the non-expertise objects. In combination, these results show that brain regions both within the face-responsive network and object- and scene-responsive networks are involved in holistic processing of faces. The neural processing involved with holistic processing of non-expertise objects are likely either weaker than those for faces (too weak for us to detect in this study) or in different brain regions, outside of the regions we localized in this study.

The shared properties that we investigated in this thesis represent only a small subsection of the many properties that are shared to varying degrees between faces, bodies and objects. Three examples of shared properties that we did not investigate in this work are emotional expression, other shared subcategories beyond those we investigated in Chapter 2 and object orientation. Emotional expression can be conveyed by both the face and the body, and thus this information could be encoded in a shared manner. Interestingly, one previous study has investigated a closely related question of whether any brain regions respond to emotional expression in an abstract manner across neural activity evoked by dynamic faces, dynamic bodies and voice stimuli (Peelen, Atkinson, & Vuilleumier, 2010). The authors found two regions showing this abstract activity one in the medial prefrontal cortex and the other in the left superior temporal sulcus. Interestingly, the coordinates of this left superior temporal sulcus region are in between the location of the body-responsive EBA and face-responsive pSTS, which could suggest a similar role of the intersection of these face- and body-responsive regions in shared face and body coding as we find at the intersection of the OFA and EBA in Chapter 4 for shared face and body orientation information.

Many subcategories are shared by faces and bodies, and we investigated the neural coding of two of these subcategories, sex and weight, in Chapter 2. We selected these two subcategories as they are easily definable subcategories that are clearly visible from both the body and face. Other examples of shared subcategories include race, age and also more abstract attributes such as attractiveness and trustworthiness. Future work will be needed to investigate how other kinds of shared subcategories are encoded.

In Chapter 4 we investigated face and body orientation, but not object orientation. A previous study found that both face and object orientation could be decoded from neural activity in the object-responsive lateral occipital (LO) area (Ramírez et al., 2014). This region is close to the region we find in Chapter 4 that encoded orientation in an abstract manner across neural activity evoked by faces and bodies. This suggests that this abstract coding could potentially also apply to objects. However, it is also possible that object orientation may be encoded differently to face and body orientation, as the view that is defined as the front view of an object can be more ambiguous than the front of a face or a body (Yangqing Xu & Franconeri, 2012).

1.6. General discussion

The Chapters of this thesis have investigated the neural coding of four different shared properties of faces and bodies or objects, namely shared subcategories (sex and weight), shared identity and shared orientation for faces and bodies, and holistic processing for faces and objects. We tested whether these shared properties were encoded in a separated manner in brain regions previously shown to be responsive to faces, bodies and objects, or if there would be overlapping or shared coding of these properties. Our results show a mixed pattern, with some aspects of the shared properties encoded in a shared or overlapping manner, and others aspects encoded in a domain-specific manner.

In several of the studies included in this thesis, we found evidence that shared properties are encoded in a distributed, overlapping and non-domain specific manner. In Chapters 2, 3 and 4, we found we could decode face and body subcategories, identities and orientations from both face- and body-responsive brain regions, demonstrating that the face- and body-responsive regions do not solely encode visual information related to their specific domain preferences. Furthermore in Chapter 5, we found neural responses related to holistic processing of faces in object- and scene-responsive brain regions, as well as in the face-responsive FFA2, demonstrating an involvement of non-domain specific regions in holistic face processing. These findings support a previous theory that neural representations of faces and objects are widely distributed and overlapping, rather than constrained to particular face- and object-responsive areas of occipitotemporal cortex (Haxby et al., 2001), and furthermore extend this theory to the visual processing of human bodies.

Beyond this overlapping coding, we also found evidence of abstract coding of shared properties, where the same pattern of activity was evoked by both faces and bodies. Several regions were found to show abstract coding, including a variety of low-level to high-level regions. In Chapter 4, we found abstract coding of face and body orientation at the intersection of the early face and body processing regions OFA and EBA and in Chapter 2, we also found evidence of abstract coding of weight in the EBA. Abstract coding of identity was found in both early areas (V1 and right inferior occipital cortex) as well as higher-level regions, the right parahippocampal cortex and right superior parietal cortex. These results show that multiple regions are involved in combining face and body information in an

abstract manner. Furthermore, they demonstrate that this information is shared throughout different levels of visual processing, and that the level the information is shared seems to depend on the specific property.

In contrast to our findings of overlapping and abstract shared coding, we found some evidence of domain-specific coding. Interestingly, this domain-specific coding seemed to relate to high-level visual transformations of the shared properties. In Chapter 2, we found that image-size invariant representations of body subcategories were mostly specific to the body-responsive brain regions. Similarly, in Chapter 3 we found viewpoint-invariant body identity representations in the body-responsive FBA, but not consistently in any face-responsive regions. These results suggest that visual transformations such as image-size and viewpoint transformations may be computed in primarily in domain-specific brain regions. These transformations are quite complex abilities of the human visual system, thus it is possible that they require dedicated processing regions, which could be one reason that we find prominent face-, body- and object-responsive brain regions in occipitotemporal cortex.

1.7. Conclusion

In this work we investigated the neural coding of a variety of shared properties of faces, bodies and objects, and tested whether they are encoded in domain-specific regions, or in an overlapping or shared neural code. We found evidence for a mix of overlapping coding, shared coding and domain-specific coding, depending on the specific property and the level of abstraction of its neural representation. These findings extend our understanding of the neural processes involved in the encoding of face, body and object properties. Furthermore, they give new insights into the overall functional organisation of occipitotemporal cortex.

1.8. Reference list

Anzellotti, S., Fairhall, S. L., & Caramazza, A. (2014). Decoding Representations of Face Identity That are Tolerant to Rotation. *Cerebral Cortex*, *24*(8), 1988–1995. https://doi.org/10.1093/cercor/bht046

Arcurio, L. R., Gold, J. M., & James, T. W. (2012). The response of face-selective cortex with single face parts and part combinations. *Neuropsychologia*, *50*(10), 2454–2459. https://doi.org/10.1016/j.neuropsychologia.2012.06.016

Axelrod, V., & Yovel, G. (2012). Hierarchical Processing of Face Viewpoint in Human Visual Cortex. *The Journal of Neuroscience*, *32*(7), 2442–2452. https://doi.org/10.1523/JNEUROSCI.4770-11.2012

Bauser, D. A. S., Suchan, B., & Daum, I. (2011). Differences between perception of human faces and body shapes: Evidence from the composite illusion. *Vision Research*, *51*(1), 195–202. https://doi.org/10.1016/j.visres.2010.11.007

Behrmann, M., & Avidan, G. (2005). Congenital prosopagnosia: Face-blind from birth. *Trends in Cognitive Sciences*, *9*(4), 180–187. https://doi.org/10.1016/j.tics.2005.02.011

Bonemei, R., Costantino, A. I., Battistel, I., & Rivolta, D. (2018). The perception of (naked only) bodies and faceless heads relies on holistic processing: Evidence from the inversion effect. *British Journal of Psychology*, *109*(2), 232–243. https://doi.org/10.1111/bjop.12271

Bracci, S., & Op de Beeck, H. (2016). Dissociations and associations between shape and category representations in the two visual pathways. *Journal of Neuroscience*, *36*(2), 432–444. https://doi.org/10.1523/JNEUROSCI.2314-15.2016

Brandman, T., & Yovel, G. (2010). The Body Inversion Effect Is Mediated by Face-Selective, Not Body-Selective, Mechanisms. *Journal of Neuroscience*, *30*(31), 10534–10540. https://doi.org/10.1523/JNEUROSCI.0911-10.2010

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, *77*, 305–327.

Carlin, J. D., & Calder, A. J. (2013). The neural basis of eye gaze processing. *Current Opinion in Neurobiology*, *23*(3), 450–455. https://doi.org/10.1016/j.conb.2012.11.014

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. https://doi.org/10.1016/j.cub.2011.09.025

Cohen, L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, M., & Michel, F. (2000). The visual word form area: spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain*, *123*, 291–307.

Curby, K. M., & Moerel, D. (2019). Behind the face of holistic perception: Holistic processing of Gestalt stimuli and faces recruit overlapping perceptual mechanisms. *Attention, Perception, and Psychophysics*, *81*(8), 2873–2880. https://doi.org/10.3758/s13414-019-01749-w

Dilks, D. D., Julian, J. B., Kubilius, J., Spelke, E. S., & Kanwisher, N. (2011). Mirror-Image Sensitivity and Invariance in Object and Scene Processing Pathways, *31*(31), 11305–11312. https://doi.org/10.1523/JNEUROSCI.1935-11.2011

Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science*, *293*(5539), 2470–2473. https://doi.org/10.1126/science.1063414

Epstein, R. (2008). Parahippocampal and retrosplenial contributions to human spatial navigation. *Trends in Cognitive Sciences*, *12*(10), 388–396. https://doi.org/10.1016/j.tics.2008.07.004

Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, *392*(6676), 598–601. https://doi.org/10.1038/33402

Flack, T. R., Harris, R. J., Young, A. W., & Andrews, T. J. (2019). Symmetrical Viewpoint Representations in Face-

Selective Regions Convey an Advantage in the Perception and Recognition of Faces. *The Journal of Neuroscience*, *39*(19), 3741–3751. https://doi.org/10.1523/jneurosci.1977-18.2019

Freiwald, W. A., & Tsao, D. Y. (2010). Functional Compartmentalization and Viewpoint Generalization Within the Macaque Face-Processing System. *Science*, *330*(6005), 845–851. https://doi.org/10.1126/science.1194908

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience*, *3*(2), 191–197. https://doi.org/10.1038/72140

Gauthier, I., & Tarr, M. J. (2002). Unraveling mechanisms for expert object recognition: bridging brain activity and behavior. *Journal of Experimental Psychology. Human Perception and Performance*, *28*(2), 431–446. https://doi.org/10.1037/0096-1523.28.2.431

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), 568–573.

Gauthier, I., Tarr, M. J., Moylan, J., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). The fusiform "face area" is part of a network that processes faces at the individual level. *Journal of Cognitive Neuroscience*, *12*(3), 495–504. https://doi.org/10.1162/089892900562165

Goffaux, V., Schiltz, C., Mur, M., & Goebel, R. (2013). Local discriminability determines the strength of holistic processing for faces in the fusiform face area. *Frontiers in Psychology*, *3*(604), 1–14. https://doi.org/10.3389/fpsyg.2012.00604

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. https://doi.org/10.1038/nn1224

Grill-Spector, K., Kushnir, T., Edelman, S., Itzchak, Y., & Malach, R. (1998). Cue-invariant activation in object-related areas of the human occipital lobe. *Neuron*, *21*(1), 191–202. https://doi.org/10.1016/S0896-6273(00)80526-7

Grill-Spector, K., Kushnir, T., Hendler, T., Edelman, S., Itzchak, Y., & Malach, R. (1998). A sequence of object processing stages revealed by fMRI in the human occipital lobe. *Human Brain Mapping*, *6*, 316–328.

Grill-Spector, K., & Weiner, K. S. (2014). The functional architecture of the ventral temporal cortex and its role in categorization. *Nature Reviews Neuroscience*, *15*, 536–548. https://doi.org/10.1038/nrn3747

Gross, C. G., Roch-Miranda, C. E., & Bender, D. B. (1972). Visual Properties of Neurons in Inferotemporal of the Macaque. *Journal of Neurophysiology*, *35*(1), 96–111.

Guntupalli, J. S., Wheeler, K. G., & Gobbini, M. I. (2017). Disentangling the Representation of Identity from Head View Along the Human Face Processing Pathway. *Cerebral Cortex*, *27*(1), 46–53. https://doi.org/10.1093/cercor/bhw344

Hahn, C. A., O'Toole, A. J., & Phillips, P. J. (2015). Dissecting the time course of person recognition in natural viewing environments. *British Journal of Psychology*, *107*(1), 117–134. https://doi.org/10.1111/bjop.12125

Halgren, E., Dale, A. M., Sereno, M. I., Tootell, R. B. H., Marinkovic, K., & Rosen, B. R. (1999). Location of human face-selective cortex with respect to retinotopic areas. *Human Brain Mapping*, *7*(1), 29–37. https://doi.org/10.1002/(SICI)1097-0193(1999)7:1<29::AID-HBM3>3.0.CO;2-R

Harris, A., Vyas, D. B., & Reed, C. L. (2016). Holistic processing for bodies and body parts: New evidence from stereoscopic depth manipulations. *Psychonomic Bulletin & Review*, *23*, 1513–1519. https://doi.org/10.3758/s13423-016-1027-4

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), 21164–21169. https://doi.org/10.1073/pnas.1212207110

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

Hasson, U., Levy, I., Behrmann, M., Hendler, T., & Malach, R. (2002). Eccentricity bias as an organizing principle for human high-order object areas. *Neuron*, *34*(3), 479–490. https://doi.org/10.1016/S0896-6273(02)00662-1

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, *293*(5539), 2425–2430. https://doi.org/10.1126/science.1063736

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*(6), 223–233. https://doi.org/10.1016/S1364-6613(00)01482-0

Haxby, J. V., Ungerleider, L. G., Clark, V. P., Schouten, J. L., Hoffman, E. A., & Martin, A. (1999). The effect of face inversion on activity in human neural systems for face and object perception. *Neuron*, *22*(1), 189–199. https://doi.org/10.1016/S0896-6273(00)80690-X

He, C., Peelen, M. V., Han, Z., Lin, N., Caramazza, A., & Bi, Y. (2013). Selectivity for large nonmanipulable objects in scene-selective visual cortex does not require visual experience. *NeuroImage*, *79*, 1–9. https://doi.org/10.1016/j.neuroimage.2013.04.051

Henriksson, L., Mur, M., & Kriegeskorte, N. (2015). Faciotopy—A face-feature map with face-like topology in the human occipital face area. *Cortex*, *72*, 156–167. https://doi.org/10.1016/j.cortex.2015.06.030

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), 80–84. https://doi.org/10.1038/71152

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, *160*(1), 106–154. https://doi.org/10.1113/jphysiol.1962.sp006837

Ishai, A. (2008). Let's face it: It's a cortical network. *NeuroImage*, *40*(2), 415–419. https://doi.org/10.1016/j.neuroimage.2007.10.040

Kamps, F. S., Julian, J. B., Kubilius, J., Kanwisher, N., & Dilks, D. D. (2016). The occipital place area represents the local elements of scenes. *NeuroImage*, *132*, 417–424. https://doi.org/10.1016/j.neuroimage.2016.02.062

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, *17*(11), 4302–4311.

Kanwisher, N., Tong, F., & Nakayama, K. (1998). The effect of face inversion on the human fusiform face area. *Cognition*, *68*(1), 1–11. https://doi.org/10.1016/S0010-0277(98)00035-3

Kietzmann, T. C., Swisher, J. D., König, P., & Tong, F. (2012). Prevalence of Selectivity for Mirror-Symmetric Views of Faces in the Ventral and Dorsal Visual Pathways. *The Journal of Neuroscience*, *32*(34), 11763–11772. https://doi.org/10.1523/jneurosci.0126-12.2012

Kitada, R., Okamoto, Y., Sasaki, A. T., Kochiyama, T., Miyahara, M., Lederman, S. J., & Sadato, N. (2013). Early visual experience and the recognition of basic facial expressions: Involvement of the middle temporal and inferior frontal gyri during haptic identification by the early blind. *Frontiers in Human Neuroscience*, *7*(7), 1–15. https://doi.org/10.3389/fnhum.2013.00007

Kitada, R., Yoshihara, K., Sasaki, A. T., Hashiguchi, M., Kochiyama, T., & Sadato, N. (2014). The brain network underlying the recognition of hand gestures in the blind: The supramodal role of the extrastriate body area. *Journal of Neuroscience*, *34*(30), 10096–10108. https://doi.org/10.1523/JNEUROSCI.0500-14.2014

Konkle, T., & Oliva, A. (2012). A Real-World Size Organization of Object Responses in Occipitotemporal Cortex. *Neuron*, *74*(6), 1114–1124. https://doi.org/10.1016/j.neuron.2012.04.036

Kriegeskorte, N., Mur, M., Ruff, D. A., Kiani, R., Bodurka, J., Esteky, H., … Bandettini, P. A. (2008). Matching Categorical Object Representations in Inferior Temporal Cortex of Man and Monkey. *Neuron*, *60*(6), 1126–1141. https://doi.org/10.1016/j.neuron.2008.10.043

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Neural Information Processing Systems.*, *25*. https://doi.org/10.1145/3065386

Kumar, S., Popivanov, I. D., & Vogels, R. (2019). Transformation of Visual Representations Across Ventral Stream Body-selective Patches. *Cerebral Cortex*, *29*(1), 215–229. https://doi.org/10.1093/cercor/bhx320

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*, 436–444. https://doi.org/10.1038/nature14539

Levy, I., Hasson, U., Avidan, G., Hendler, T., & Malach, R. (2001). Center-periphery organization of human object areas. *Nature Neuroscience*, *4*(5), 533–539. https://doi.org/10.1038/87490

Liu, J., Harris, A., & Kanwisher, N. (2010). Perception of face parts and face configurations: an fMRI study. *Journal of Cognitive Neuroscience*, *22*(1), 203–211. https://doi.org/10.1162/jocn.2009.21203

Loffler, G., Yourganov, G., Wilkinson, F., & Wilson, H. R. (2005). fMRI evidence for the neural representation of faces. *Nature Neuroscience*, *8*(10), 1386–1390. https://doi.org/10.1038/nn1538

Logothetis, N. K. (2002). The neural basis of the blood-oxygen-level-dependent functional magnetic resonance imaging signal. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *357*(1424), 1003–1037. https://doi.org/10.1098/rstb.2002.1114

Mahon, B. Z., Anzellotti, S., Schwarzbach, J., Zampini, M., & Caramazza, A. (2009). Category-Specific Organization in the Human Brain Does Not Require Visual Experience. *Neuron*, *63*(3), 397–405. https://doi.org/10.1016/j.neuron.2009.07.012

Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., … Tootell, R. B. H. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(18), 8135–8139. https://doi.org/10.1073/pnas.92.18.8135

Masland, R. H. (2012). The Neuronal Organization of the Retina. *Neuron*, *76*(2), 266–280. https://doi.org/10.1016/j.neuron.2012.10.002

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

McCandliss, B. D., Cohen, L., & Dehaene, S. (2003). The visual word form area: Expertise for reading in the fusiform gyrus. *Trends in Cognitive Sciences*, *7*(7), 293–299. https://doi.org/10.1016/S1364-6613(03)00134-7

Nasr, S., & Tootell, R. B. H. (2012). Role of fusiform and anterior temporal cortical areas in facial recognition. *NeuroImage*, *63*(3), 1743–1753. https://doi.org/10.1016/j.neuroimage.2012.08.031

O'Toole, A. J., Phillips, P. J., Weimer, S., Roark, D. A., Ayyad, J., Barwick, R., & Dunlop, J. (2011). Recognizing people from dynamic and static faces and bodies: Dissecting identity with a fusion approach. *Vision Research*, *51*(1), 74–83. https://doi.org/10.1016/j.visres.2010.09.035

Ogawa, S., Lee, T. -M, Nayak, A. S., & Glynn, P. (1990). Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields. *Magnetic Resonance in Medicine*, *14*(1), 68–78. https://doi.org/10.1002/mrm.1910140108

Ogawa, S., Lee, T. M., Kay, A. R., & Tank, D. W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences*, *87*(24), 9868–9872. https://doi.org/10.1073/pnas.87.24.9868

Op de Beeck, H. P., Brants, M., Baeck, A., & Wagemans, J. (2010). Distributed subordinate specificity for bodies, faces, and buildings in human ventral visual cortex. *NeuroImage*, *49*(4), 3414–3425. https://doi.org/10.1016/j.neuroimage.2009.11.022

Op de Beeck, H. P., Pillet, I., & Ritchie, J. B. (2019). Factors Determining Where Category-Selective Areas Emerge in Visual Cortex. *Trends in Cognitive Sciences*, *23*(9), 784–797. https://doi.org/10.1016/j.tics.2019.06.006

Orlov, T., Makin, T. R., & Zohary, E. (2010). Topographic Representation of the Human Body in the Occipitotemporal Cortex. *Neuron*, *68*(3), 586–600. https://doi.org/10.1016/j.neuron.2010.09.032

Peelen, M. V., Atkinson, A. P., & Vuilleumier, P. (2010). Supramodal Representations of Perceived Emotions in the Human Brain. *Journal of Neuroscience*, *30*(30), 10127–10134. https://doi.org/10.1523/JNEUROSCI.2161-10.2010

Peelen, M. V, & Downing, P. E. (2005). Selectivity for the human body in the fusiform gyrus. *Journal of Neurophysiology*, *93*(1), 603–608. https://doi.org/10.1152/jn.00513.2004

Pietrini, P., Furey, M. L., Ricciardi, E., Gobbini, M. I., Wu, W.-H. C., Cohen, L., … Haxby, J. V. (2004). Beyond sensory images: Object-based representation in the human ventral pathway. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(15), 5658–5663. https://doi.org/10.1073/pnas.0400707101

Premereur, E., Taubert, J., Janssen, P., Vogels, R., & Vanduffel, W. (2016). Effective Connectivity Reveals Largely Independent Parallel Networks of Face and Body Patches. *Current Biology*, *26*(24), 3269–3279. https://doi.org/10.1016/j.cub.2016.09.059

Rajimehr, R., Young, J. C., & Tootell, R. B. H. (2009). An anterior temporal face patch in human cortex, predicted by macaque maps. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(6), 1995–2000. https://doi.org/10.1073/pnas.0807304106

Ramírez, F. M. (2018). Orientation Encoding and Viewpoint Invariance in Face Recognition: Inferring Neural Properties from Large-Scale Signals. *The Neuroscientist*, *24*(6), 582–608. https://doi.org/10.1177/1073858418769554

Ramírez, F. M., Cichy, R. M., Allefeld, C., & Haynes, J.-D. (2014). The neural code for face orientation in the human fusiform face area. *The Journal of Neuroscience*, *34*(36), 12155–12167. https://doi.org/10.1523/JNEUROSCI.3156-13.2014

Reich, L., Szwed, M., Cohen, L., & Amedi, A. (2011). A ventral visual stream reading center independent of visual experience. *Current Biology*, *21*(5), 363–368. https://doi.org/10.1016/j.cub.2011.01.040

Rice, A., Phillips, P. J., Natu, V., An, X., & O'Toole, A. J. (2013). Unaware Person Recognition From the Body When Face Identification Fails. *Psychological Science*, *24*(11), 2235–2243. https://doi.org/10.1177/0956797613492986

Rice, A., Phillips, P. J., & O'Toole, A. (2013). The role of the face and body in unfamiliar person identification. *Applied Cognitive Psychology*, *27*(6), 761–768. https://doi.org/10.1002/acp.2969

Richler, J. J., & Gauthier, I. (2014). A meta-analysis and review of holistic face processing. *Psychological Bulletin*, *140*(5), 1281–1302. https://doi.org/10.1037/a0037004

Robbins, R. A., & Coltheart, M. (2012a). Left–right holistic integration of human bodies. *The Quarterly Journal of Experimental Psychology*, *65*(10), 1962–1974. https://doi.org/10.1080/17470218.2012.674145

Robbins, R. A., & Coltheart, M. (2012b). The effects of inversion and familiarity on face versus body cues to person recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(5), 1098–1104. https://doi.org/10.1037/a0028584

Ross, D. A., Tamber-Rosenau, B. J., Palermi, T. J., Zhang, J., Xu, Y., & Gauthier, I. (2018). High-resolution Functional Magnetic Resonance Imaging Reveals Configural Processing of Cars in Right Anterior Fusiform Face Area of Car Experts. *Journal of Cognitive Neuroscience*, *30*(7), 973–984. https://doi.org/10.1162/jocn_a_01256

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.

https://doi.org/10.1038/nn1370

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), 1–16. https://doi.org/10.1167/10.2.25

Schiltz, C., & Rossion, B. (2006). Faces are represented holistically in the human occipito-temporal cortex. *NeuroImage*, *32*(3), 1385–1394. https://doi.org/10.1016/j.neuroimage.2006.05.037

Schultz, J., & Pilz, K. S. (2009). Natural facial motion enhances cortical responses to faces. *Experimental Brain Research*, *194*(3), 465–475. https://doi.org/10.1007/s00221-009-1721-9

Schwarzlose, R. F., Baker, C. I., & Kanwisher, N. (2005). Separate Face and Body Selectivity on the Fusiform Gyrus. *The Journal of Neuroscience*, *25*(47), 11055–11059. https://doi.org/10.1523/JNEUROSCI.2621-05.2005

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. https://doi.org/10.1523/JNEUROSCI.1704-15.2016

Szegedy, C., Zaremba, W., Sutskever, I., Bruna, J., Erhan, D., Goodfellow, I., & Fergus, R. (2014). Intriguing properties of neural networks. *International Conference on Learning Representations*, 1–10.

Taylor, J. C., Wiggett, A. J., & Downing, P. E. (2007). Functional MRI analysis of body and body part representations in the extrastriate and fusiform body areas. *Journal of Neurophysiology*, *98*(3), 1626–1633. https://doi.org/10.1152/jn.00012.2007

Tong, F., Nakayama, K., Moscovitch, M., Weinrib, O., & Kanwisher, N. (2000). Response properties of the human fusiform face area. *Cognitive Neuropsychology*, *17*(1–3), 257–279. https://doi.org/10.1080/026432900380607

Tsao, D. Y., Moeller, S., & Freiwald, W. A. (2008). Comparing face patch systems in macaques and humans. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(49), 19514–19519. https://doi.org/10.1073/pnas.0809662105

van den Hurk, J., Van Baelen, M., & Op de Beeck, H. P. (2017). Development of visual category selectivity in ventral visual cortex does not require visual experience. *Proceedings of the National Academy of Sciences*, *114*(22), E4501–E4510. https://doi.org/10.1073/pnas.1612862114

Wandell, B. A., Dumoulin, S. O., & Brewer, A. A. (2007). Visual field maps in human cortex. *Neuron*, *56*(2), 366–383. https://doi.org/10.1016/j.neuron.2007.10.012

Willems, S., Vrancken, L., Germeys, F., & Verfaillie, K. (2014). Holistic processing of human body postures: Evidence from the composite effect. *Frontiers in Psychology*, *5*(618), 1–9. https://doi.org/10.3389/fpsyg.2014.00618

Wong, A. C.-N., Palmeri, T. J., Rogers, B. P., Gore, J. C., & Gauthier, I. (2009). Beyond shape: How you learn about objects affects how they are represented in visual cortex. *PLoS ONE*, *4*(12), 1–7. https://doi.org/10.1371/journal.pone.0008405

Xu, Yangqing, & Franconeri, S. L. (2012). The Head of the Table: Marking the “Front” of An Object Is Tightly Linked with Selection. *Journal of Neuroscience*, *32*(4), 1408–1412. https://doi.org/10.1523/JNEUROSCI.4185-11.2012

Xu, Yaoda. (2005). Revisiting the role of the fusiform face area in visual expertise. *Cerebral Cortex*, *15*(8), 1234–1242. https://doi.org/10.1093/cercor/bhi006

Yarbus, A. L. (1967). *Eye movements and vision*. New York: Plenum Press.

Yovel, G., Pelc, T., & Lubetzky, I. (2010). It’s all in your head: why is the body inversion effect abolished for headless bodies? *Journal of Experimental Psychology. Human Perception and Performance*, *36*(3), 759–767. https://doi.org/10.1037/a0017451

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. https://doi.org/10.1016/j.neuroimage.2016.01.045

Zhao, M., Bülthoff, H. H., & Bülthoff, I. (2016). Beyond faces and expertise: Face-like holistic processing of nonface objects in the absence of expertise. *Psychological Science*, *27*(2), 213–222. https://doi.org/10.1177/0956797615617779

1.9. Declaration of Contribution

This doctoral thesis comprises of five manuscripts that are either accepted or in preparation for publication. Details of the contribution of the candidate and co-authors to these manuscripts are as follows:

1. Foster, C., Zhao, M., Romero, J., Black, M. J., Mohler, B. J., Bartels, A. & Bülthoff, I. (2019) Decoding subcategories of human bodies from both body- and face-responsive cortical regions. *NeuroImage, 202*, 1–13.

   C.F., M.Z., A.B., I.B. and B.J.M. designed the study, J.R. and M.J.B. assisted with stimuli generation, C.F. programmed the experiment, collected and analysed the data and wrote the first version of the manuscript, all authors contributed to revisions of the manuscript.
2. Foster, C., Zhao, M., Bolkart, T., Black, M. J., Bartels, A. & Bülthoff, I. (in preparation) The neural coding of face and body identity.

   C.F., M.Z., A.B. and I.B. designed the study, T.B. and M.J.B. assisted with stimuli generation, C.F. programmed the experiment, collected and analysed the data and wrote the first version of the manuscript, all authors contributed to revisions of the manuscript.
3. Foster, C., Zhao, M., Bolkart, T., Black, M. J., Bartels, A. & Bülthoff, I. (in preparation) An abstract neural code in occipital cortex for person orientation.

   C.F., M.Z., A.B. and I.B. designed the study, T.B. and M.J.B. assisted with stimuli generation, C.F. programmed the experiment, collected and analysed the data and wrote the first version of the manuscript, all authors contributed to revisions of the manuscript.
4. Foster, C., Bülthoff, I. Bartels, A. & Zhao, M., (in preparation)
   Investigating holistic face processing within and outside of face-responsive brain regions.

   C.F., I.B., A.B. and M.Z. designed the study, C.F. programmed the experiment, collected and analysed the data and wrote the first version of the manuscript, all authors contributed to revisions of the manuscript.

5. Foster, C., Bülthoff, I. Bartels, A. & Zhao, M., (in preparation)
   No holistic processing of objects in brain regions that process faces holistically, despite an identical behavioural effect.
   C.F., I.B., A.B. and M.Z. designed the study, C.F. programmed the experiment, collected and analysed the data and wrote the first version of the manuscript, all authors contributed to revisions of the manuscript.

Parts of this work were also presented at the following conferences:

1. Foster, C., Zhao, M., Romero, J., Black, M.J., Mohler, B.J., Bartels, A. & Bülthoff, I. (August 2017) Decoding categories shared by the face and body. 40th European Conference on Visual Perception (ECVP 2017), Berlin, Germany.
2. Foster, C., Zhao, M., Romero, J., Black, M.J., Mohler, B.J., Bartels, A. & Bülthoff, I. (October 2017) Decoding categories shared by the face and body. 18th Conference of Junior Neuroscientists of Tübingen (NeNa 2017), Schramberg, Germany.
3. Foster, C., Zhao, M., Bartels, A. & Bülthoff, I (May 2018) Neural Correlates of Holistic Face Processing, 18th Annual Meeting of the Vision Sciences Society (VSS 2018), St. Pete Beach, FL, USA.
4. Foster, C., Zhao, M., Bartels, A. and Bülthoff, I. (August 2018) No holistic processing of objects in brain regions that process faces holistically, despite an identical behavioural effect, 41st European Conference on Visual Perception (ECVP 2018), Trieste, Italy.
5. Foster, C., Zhao, M., Bolkart, T., Black, M.J., Bartels, A. & Bülthoff, I. (May 2019) Decoding the viewpoint and identity of faces and bodies. 19th Annual Meeting of the Vision Sciences Society (VSS 2019), St. Pete Beach, FL, USA.

2. Decoding subcategories of human bodies from both body- and face-responsive cortical regions

NeuroImage 202 (2019) 116085

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Decoding subcategories of human bodies from both body- and face-responsive cortical regions

Celia Foster [a,b,c,*], Mintao Zhao [a,d], Javier Romero [b], Michael J. Black [b], Betty J. Mohler [a,b], Andreas Bartels [a,c,e,f], Isabelle Bülthoff [a,**]

[a] Max Planck Institute for Biological Cybernetics, Tübingen, Germany
[b] Max Planck Institute for Intelligent Systems, Tübingen, Germany
[c] Centre for Integrative Neuroscience, Tübingen, Germany
[d] School of Psychology, University of East Anglia, UK
[e] Department of Psychology, University of Tübingen, Germany
[f] Bernstein Center for Computational Neuroscience, Tübingen, Germany

ARTICLE INFO

Keywords:
Body perception
Face perception
EBA
FBA
OFA
FFA

ABSTRACT

Our visual system can easily categorize objects (e.g. faces vs. bodies) and further differentiate them into subcategories (e.g. male vs. female). This ability is particularly important for objects of social significance, such as human faces and bodies. While many studies have demonstrated category selectivity to faces and bodies in the brain, how subcategories of faces and bodies are represented remains unclear. Here, we investigated how the brain encodes two prominent subcategories shared by both faces and bodies, sex and weight, and whether neural responses to these subcategories rely on low-level visual, high-level visual or semantic similarity. We recorded brain activity with fMRI while participants viewed faces and bodies that varied in sex, weight, and image size. The results showed that the sex of bodies can be decoded from both body- and face-responsive brain areas, with the former exhibiting more consistent size-invariant decoding than the latter. Body weight could also be decoded in face-responsive areas and in distributed body-responsive areas, and this decoding was also invariant to image size. The weight of faces could be decoded from the fusiform body area (FBA), and weight could be decoded across face and body stimuli in the extrastriate body area (EBA) and a distributed body-responsive area. The sex of well-controlled faces (e.g. excluding hairstyles) could not be decoded from face- or body-responsive regions. These results demonstrate that both face- and body-responsive brain regions encode information that can distinguish the sex and weight of bodies. Moreover, the neural patterns corresponding to sex and weight were invariant to image size and could sometimes generalize across face and body stimuli, suggesting that such subcategorical information is encoded with a high-level visual or semantic code.

1. Introduction

Our visual system makes use of various aspects of shape information to categorize objects (e.g. person vs. house) and to further categorize them into different subcategories (e.g. male vs. female person). This seemingly effortless ability is actually remarkably non-trivial, as objects that fit one subcategory can be of a great variability (e.g. both faces and bodies can belong to the same subcategory male), yet exemplars from different subcategories can look comparably similar (e.g. male vs. female faces). In the brain, both monkey neurophysiology and human neuroimaging studies indicate that object categorization and subcategorization processes are primarily implemented in the ventral temporal cortex (VTC) (Grill-Spector and Weiner, 2014; Gross et al., 1972; Haxby et al., 2001; Kriegeskorte et al., 2008; Logothetis and Sheinberg, 1996; Tanaka, 1996). While the VTC contains high-level category-selective areas for objects (Malach et al., 1995), faces (Gauthier et al., 2000; Kanwisher et al., 1997), bodies (Downing et al., 2001; Peelen and Downing, 2005), scenes (Epstein and Kanwisher, 1998), and visually presented words (Cohen et al., 2000), how, and where in the brain, subcategories are encoded remains to be elucidated. Even less is known about how the brain represents the same semantic categories that are shared by different object categories (e.g. sex of faces and sex of bodies).

* Corresponding author. Max Planck Institute for Biological Cybernetics, Max-Planck-Ring 8, 72076, Tübingen, Germany.
** Corresponding author.
E-mail addresses: celia.foster@tuebingen.mpg.de (C. Foster), isabelle.buelthoff@tuebingen.mpg.de (I. Bülthoff).

https://doi.org/10.1016/j.neuroimage.2019.116085
Received 20 July 2018; Received in revised form 17 July 2019; Accepted 7 August 2019
Available online 8 August 2019

Recent studies on face perception suggest that face-responsive brain areas may contain neural representations of subcategories (e.g. the sex, race, and identity of faces). Different patterns of neural responses to male and female faces have been identified in the fusiform face area (FFA) and other face-responsive regions (Contreras et al., 2013; Freeman et al., 2010; Kaul et al., 2011). Different patterns of neural responses to faces of different races have also been identified in the fusiform gyrus and early visual cortex (Contreras et al., 2013; Ratner et al., 2013). Several studies have investigated how the brain represents face identities (an extreme level of subcategorization) and found stimulus-independent representation of face identities in the anterior temporal face area (ATFA) (Anzellotti et al., 2014; Guntupalli et al., 2016; Kriegeskorte et al., 2007). Other studies found that the FFA (Anzellotti et al., 2014; Axelrod and Yovel, 2015) and the superior intraparietal sulcus (Jeong and Xu, 2016) also encode face identity.

It remains unclear exactly which subcategory features drive distinctive patterns of neural responses to different face subcategories, and whether these subcategories are represented in brain areas beyond those selective for faces (Haxby et al., 2001). The distinction between visual and semantic representation has been observed during general object categorization (Bracci and Op de Beeck, 2016), which may similarly apply to face and body categorization. For instance, different patterns of neural responses to male versus female faces could be driven by differential sensitivity to the *visual feature* of hairstyle, rather than the perceived *semantic category* of biological sex. Similarly, separable neural responses to faces of different races may be caused by the visual feature of skin tone, rather than the semantic category of race. Studies finding different neural responses to bodies of different subcategories in body-responsive areas may have the same visual/semantic concern. There is some evidence that the extrastriate body area (EBA) and fusiform body area (FBA) contain information about the identity of bodies (Ewbank et al., 2011) and the FBA and right middle occipital gyrus contain information about the weight of bodies (Hummel et al., 2013). However, different neural responses to lower vs. higher weight bodies might be supported by different sensitivity to the physical image size of bodies rather than the semantic perception of body weight. In behaviour, perceived body weight has been found to be processed independently of physical image size (Sturman et al., 2017).

In this study, we investigated how information about subcategories shared by faces and bodies is encoded in the face- and body-responsive brain networks. We chose the same two subcategories of faces and bodies: sex (male vs. female) and weight (lower vs. higher). We presented participants with images of faces and bodies varying in sex, weight, and image size (larger vs. smaller) whilst recording their brain activity using functional magnetic resonance imaging (fMRI). Inclusion of both face and body stimuli allowed us to investigate whether the brain encodes shared semantic subcategories (e.g. male/female) in an abstract manner, despite dramatic differences in the visual appearance of stimuli (e.g. faces vs. bodies). We varied image size of stimuli to further test whether the neural coding of subcategories is more abstract (i.e. image size invariant) or more bound to visual features (i.e. image size dependent). Varying image size also helps to clarify whether or not neural processes in high-level visual areas are modulated by low-level visual features (e.g. image size) (Andrews and Ewbank, 2004; Sawamura et al., 2005; Yue et al., 2011).

To investigate how face- and body-responsive brain regions encode the subcategories of sex and weight, we first trained support vector machine (SVM) classifiers to discriminate patterns of blood-oxygen-level dependent (BOLD) activity elicited by each subcategory, and then used them to predict the subcategories within separate test data. We further tested if neural responses to different subcategories generalize across stimuli sizes (e.g. trained with smaller faces, tested on larger faces), and across face and body stimuli (e.g. trained with faces, tested on bodies). We performed these multivoxel pattern analyses (MVPA) for both functionally defined face- and body-responsive areas and in whole-brain searchlight analyses. These analyses allow us to differentiate whether the neural coding of a subcategory is driven by low-level visual features, by high-level visual features, or by semantic processing. Specifically, if neural responses to face and body subcategories were driven by low-level visual features, they would be dependent on image size. If the neural responses to these subcategories were driven by high-level visual features, they would be independent of the image size, but not necessarily able to generalize across face and body stimuli. In contrast, if neural responses were driven by semantic information then they would be able to generalize across image size and across face and body stimuli.

2. Materials and methods

2.1. Participants

Thirteen participants (7 female, 6 male, 22–32 years old) were included in the fMRI experiment analyses presented in the results section, out of fifteen who had originally participated. Due to scanner malfunction the face-stimuli runs could not be completed in one participant. This participant was included in the body-related analyses but not the face ones. Data from two participants were excluded from analyses due to excessive head movement during scanning. All participants provided written informed consent prior to the experiment, and the procedures were approved by the ethics committee of the University Clinic Tübingen.

2.2. Stimuli

The experimental stimuli were grayscale images of faces and bodies that varied in biological sex (i.e. male or female) and weight (i.e. higher weight or lower weight), resulting in four stimulus classes for faces, and four for bodies. Each class was presented in both a smaller and larger image size (larger images were twice the height and width of smaller images) resulting in a total of 16 conditions (Fig. 1A). Each class contained 42 exemplars (e.g. low-weight, male bodies), all of which were different individuals. The perceived sex and weight of these images were validated via ratings from independent observers (see sections 2.2.3 and 3.1.1 below).

2.2.1. Face stimuli

Face stimuli were created using 3D face models of the face database of the Max Planck Institute for Biological Cybernetics (Blanz and Vetter, 1999; Troje and Bülthoff, 1996). To increase the variability of weight of faces, we intensified the perceived weight of the faces using a 3D morphable model (Blanz and Vetter, 1999), using the following procedure. Six observers (5 female, 1 male, 22–27 years old, 1 author), who did not participate in the fMRI experiment, rated the perceived weight of 311 faces from the database. Based on these ratings we selected all faces that were above or below one standard deviation from the average perceived weight, separately for male and female faces. These selected faces were then morphed together to generate average higher and lower weight faces separately for male and female faces. The difference between higher- and lower-weight average morphs was then applied to the original individual face models, so that for each face we were able to create higher and lower weight versions of that face. Any stimuli with artefacts from the morphing procedure were removed.

2.2.2. Body stimuli

Body stimuli were created based on body scans from the CAESAR dataset (Robinette et al., 2002). We selected bodies with a Body Mass Index (BMI) between one and two standard deviations above the average BMI for our higher weight bodies, and between one and two standard deviations below the average BMI for our lower weight bodies (both calculated separately for male and female bodies). The selected 264 body scans were then registered to a 3D body shape and pose model (Loper et al., 2015). This allowed us to obtain individual body shapes in a standard A-pose (see Fig. 1A). We modified the texture obtained from the

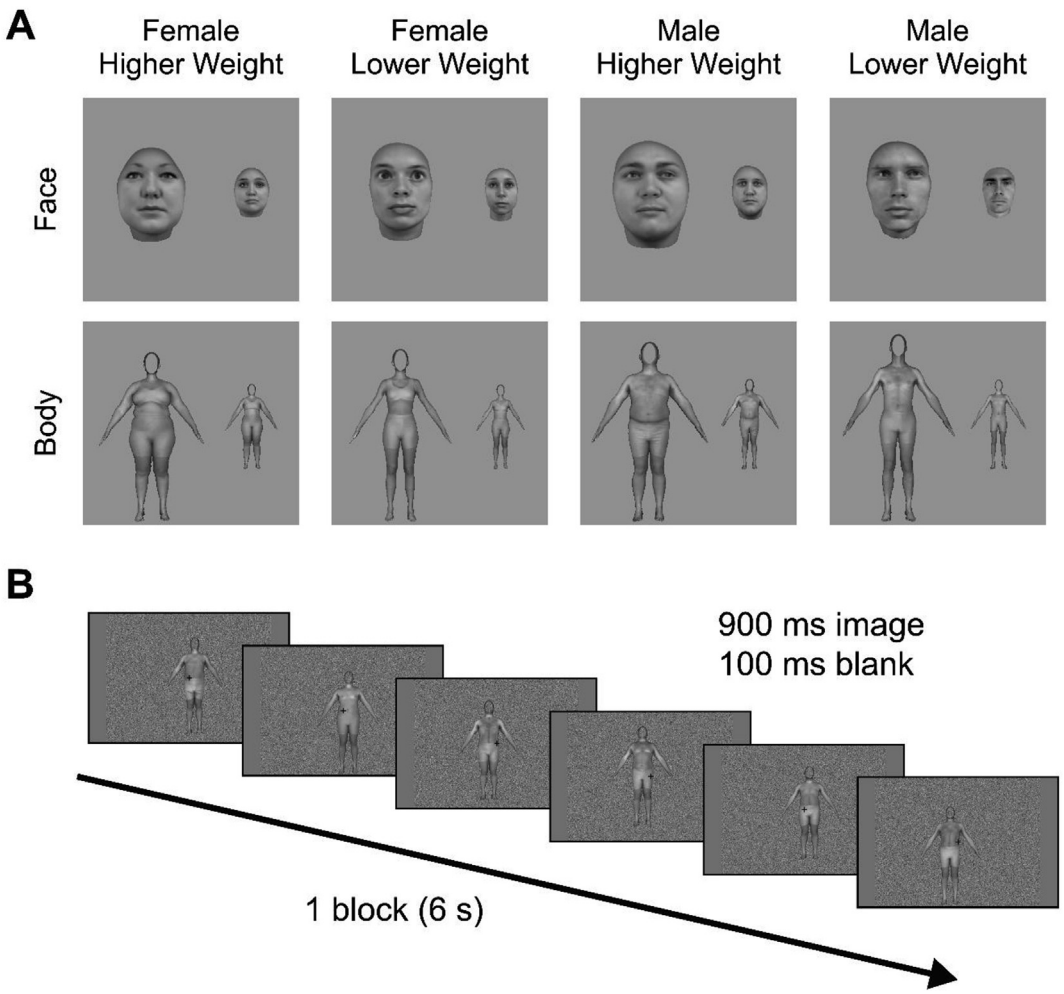

Fig. 1. Experimental stimuli and procedure of the fMRI experiment. (A) Example stimuli for the 16 conditions of the fMRI experiment. Stimuli were shown at a larger and a smaller image size (larger images were twice the height and width of smaller images). (B) An example block of stimuli in the fMRI experiment. Subjects viewed the stimuli in 6 s blocks, where each block contained images from one condition. Each image was presented for 900 m s with a 100 m s blank grayscale screen between images. Two Gaussian noise only images were shown between blocks, each lasted 900 m s followed by a 100 m s blank screen.

original scans in order to remove markers and fill in any missing texture. In the final body images, the faces were covered with an oval in order to exclude any face information from the body images.

2.2.3. Ratings of face and body stimuli

To select stimuli that truly differed in perceived weight and sex, we had six observers (3 female, 3 male, 22–35 years old), who were not participants in the fMRI experiment, rate the perceived sex and weight, for both face and body stimuli, on 7-point Likert scales. Based on these ratings we then selected 42 stimuli for each of the eight conditions (e.g. low-weight, male bodies) that maximised perceived difference in sex and weight.

2.2.4. Background stimuli

During the fMRI experiment, face and body images were shown in front of a randomly generated Gaussian noise background (Fig. 1B), in order to keep the area of retinal stimulation constant for all stimuli despite differences in the foreground image shape.

2.2.5. Localizer stimuli

Stimuli for the localizer experiment consisted of grayscale images of faces, headless bodies, objects and phase-scrambled images. Phase-scrambled images were created by making a collage containing the face and headless body images and then generating Fourier-scrambled images from the collage image.

2.3. fMRI experiment

The study consisted of two fMRI sessions on separate days. On the first day localizer runs and anatomical data were collected. On the second day experimental runs were collected. Stimuli were presented via a projector (resolution 1920×1080) with Matlab 2013b using the Psychophysics Toolbox extensions (Brainard, 1997; Kleiner et al., 2007) on a Windows PC. Participants lay supine in the scanner and viewed a screen positioned behind their head via a mirror attached to the head coil. The screen was positioned at a distance of 82 cm, and spanned $28^\circ \times 16^\circ$ of visual angle in horizontal and vertical directions respectively.

2.3.1. Experimental runs

Participants completed 8 experimental runs where 4 runs contained face stimuli and 4 runs contained body stimuli. Each run contained 8 conditions of a 2 (Sex: male vs. female) x 2 (Weight: lower vs. higher) x 2 (Image Size: larger vs. smaller) factorial design (Fig. 1A). Conditions were presented in a carryover counterbalanced blocked design, such that each condition block was preceded by each condition block once in a run (Brooks, 2012). This was to avoid biases due to remaining BOLD activation from a previous condition block (Aguirre, 2007). Stimuli were presented in front of a centred Gaussian noise background (width 5.1°,

height 7.9°). Average stimulus sizes were of the following widths and heights; larger faces 2.4° × 3.7°; smaller faces 1.2° × 1.8°; larger bodies 4.0° × 6.5°; smaller bodies 2.0° × 3.3°. In each block 6 images were shown, where each image was shown for 0.9 s followed by a 0.1 s blank grayscale screen. Two Gaussian noise only images were shown between blocks, each for 0.9 s followed by a 0.1 s blank grayscale screen. Images in each block appeared in 6 different positions (horizontally either 0.2° left or right of the screen centre; vertically either centred, or 0.2° above or below the screen centre) in a random order.

2.3.1.1. fMRI task. Participants fixated a central fixation cross at all times and pressed a button whenever a red dot appeared on the fixation cross. This occurred twice during each block at randomly determined times at least 2 s apart.

2.3.2. Localizer runs

Participants completed 5 localizer runs, which were collected on a separate day prior to the experimental runs. In each run participants viewed blocks of faces, bodies, objects and phase-scrambled images. Faces, bodies and objects were shown in front of the phase-scrambled images to match the amount of retinal stimulation in all blocks. In each block 8 images were shown, with each image being presented for 1.8 s, and a 0.2 s blank grayscale screen between images. Images were presented in a carryover counterbalanced sequence (Brooks, 2012). Participants performed a one-back matching task on the images, to ensure balanced attention across conditions. Image repetitions occurred on average once every 9 s.

2.4. fMRI scan parameters

Images were acquired using a 3T Siemens Prisma scanner with a 64-channel head coil (Siemens, Erlangen, Germany). Functional T2* echo-planar images (EPI) were acquired using a sequence with the following parameters; multiband acceleration factor 2, TR 1.2 s, TE 30 m s, flip angle 68°, FOV 192 × 192 mm. Volumes consisted of 36 slices, with an isotropic voxel size of 3 × 3 × 3 mm. The first 8 vol of each run were discarded to allow for equilibration of the T1 signal. During each session a gradient echo field map was recorded so that magnetic field inhomogeneity could be corrected during preprocessing. For each participant a high-resolution T1-weighted anatomical scan was acquired with the following parameters; TR 2 s, TE 3.06 m s, FOV 232 × 256 mm, 192 slices, isotropic voxel size of 1 × 1 × 1 mm.

2.5. fMRI data preprocessing

Data was preprocessed using SPM12 (http://www.fil.ion.ucl.ac.uk/spm/). All functional data was realigned, unwarped to correct for inhomogenities in the magnetic field and coregistered to the anatomical data. Localizer data was spatially smoothed with a 6 mm Gaussian kernel. For the whole-brain univariate analyses the data was normalized to MNI (Montreal Neurological Institute) space and spatially smoothed with a 6 mm Gaussian kernel. Multivariate data analyses were performed in individual subject space on unsmoothed data. For searchlight analyses the resulting single-subject maps of classification accuracies were normalized to MNI space, and spatially smoothed with a 6 mm Gaussian kernel.

2.6. Definition of regions of interest

We defined separate face- and body-responsive regions of interest (ROIs) using data from the localizer runs. The contrast faces > objects was used to identify the OFA and FFA and the contrast bodies > objects was used to identify the EBA and FBA (Table 1). As ROI size has been shown to affect decoding accuracy (Gardumi et al., 2016) we kept ROI sizes constant by selecting the 100 most active voxels in each region bilaterally to form the ROI. To achieve this, we initially attempted to identify each ROI using a threshold of $p < 0.05$ (FWE corrected). If the ROI was not identifiable in 100 voxels we attempted to define the ROI using a lower threshold of $t = 2$. ROIs were defined using localizer data only, therefore their definition was independent to the main experiment analyses. The FFA and the FBA are known to partially overlap (Schwarzlose et al., 2005). We initially performed analyses in both ROIs without removing any overlapping voxels (this overlap had a mean of 32% of FFA voxels and 38% of FBA voxels). In any instances where an analysis was significant for both FFA and FBA (and therefore significant decoding could be caused be overlapping voxels) we planned to run additional analyses removing the overlapping voxels to investigate this possibility.

In addition to the separate ROIs we defined distributed face- and body-responsive ROIs that contained voxels from several of the isolated ROIs, as described previously (Hahn and O'Toole, 2017). This allowed us to investigate whether neural information from distributed brain regions improves categorical classification of faces and bodies. Distributed face-responsive ROIs were defined using voxels from OFA, FFA, STS and ATFA. Distributed body-responsive ROIs were defined using voxels from EBA and FBA. The 300 and 500 most responsive voxels were selected to create two sizes of distributed ROIs. We defined 300 voxel ROIs as every participant had at least 300 face- and body-responsive voxels, thus we were able to define a distributed face- and body-responsive ROI of this size in every participant. Most participants also had many more face-and body-responsive voxels, thus we additionally defined 500 voxel face- and body-responsive ROIs (in N = 10 and N = 11 participants respectively) to see if classifier performance would benefit from information in these additional voxels.

We used V1 as a control ROI, which was bilaterally localized using anatomical labels generated using the Freesurfer software package (Hinds et al., 2009) (https://surfer.nmr.mgh.harvard.edu/). This method generates V1 ROIs based on the anatomy of the participant, and the method has been validated to show that there is close agreement between these anatomical V1 ROIs and V1 defined functionally using retinotopic mapping. For each participant, we selected the 50 most posterior voxels (corresponding to the foveal section of V1) from each hemisphere and combined them to create a 100-voxel V1 ROI.

Table 1
ROI coordinates (in MNI space), ROI volume and number of subjects each ROI was identified in (N). All ROI analyses were done in subject space. ROIs were subsequently normalized to MNI space in order to show the mean ROI locations. Coordinates show mean x, y and z locations and volume, ± standard deviations.

ROI	hemisphere	x	y	z	Volume (mm^3)	N
OFA	left	−34 ± 6.4	−83 ± 7.1	−12 ± 3.4	520 ± 296.9	13
	right	37 ± 5.9	−80 ± 7.4	−13 ± 3.5	945 ± 356.4	13
FFA	left	−41 ± 2.8	−56 ± 8.8	−19 ± 3.7	519 ± 253.4	12
	right	43 ± 3.1	−53 ± 6.0	−18 ± 2.6	979 ± 274.5	13
EBA	left	−47 ± 2.8	−77 ± 3.7	4 ± 5.0	385 ± 153.0	12
	right	48 ± 2.6	−70 ± 3.3	−1 ± 5.4	1106 ± 252.1	13
FBA	left	−42 ± 2.7	−48 ± 5.8	−21 ± 5.0	374 ± 277.2	11
	right	44 ± 2.9	−49 ± 4.0	−18 ± 3.5	994 ± 272.2	11
V1	left	−9 ± 1.4	−100 ± 1.8	−3 ± 4.3	390 ± 70.8	13
	right	12 ± 3.9	−98 ± 1.8	0 ± 3.9	468 ± 80.2	13

To investigate if our choice to fix the ROI size at a constant number of voxels, rather than fixing ROI size at a constant t-contrast threshold had any impact on our results, we additionally conducted supplemental analyses where ROIs were defined using this alternative method (Supplementary Table S1). Specifically, we defined supplemental versions of our face- and body-responsive ROIs by selecting all active voxels at a constant threshold of $t = 3$ rather than a constant number of voxels. A supplemental V1 ROI was defined using the same method as the original V1 ROI, except that the 510 most posterior voxels were selected in order to match the size of this V1 ROI to the average size of the largest separate ROI (EBA). All results of these supplemental analyses can be found in the supplemental materials.

2.7. Multivoxel pattern analyses (MVPA)

Data from the main experiment was modelled with General Linear Models (GLMs) using SPM12. The responses to each block (i.e. trial) were modelled as separate regressors in the GLM. Classification analyses were performed using The Decoding Toolbox (Hebart et al., 2015) from the resulting beta weight images. Input data was feature-scaled using z-score normalization and outlier values (greater than 2 standard deviations) were set to 2 or -2. Mean and standard deviation for z-scoring were estimated using training data and then these values were applied to testing data to ensure independence of training and testing data. Classification was performed using a linear support vector machine classifier (LIBSVM).

We performed three different MVPA analyses (one combining both image sizes, one across image size, and one across face and body stimuli), and each was done for separate ROIs, distributed ROIs and in whole-brain searchlight analyses. Statistical significance for ROI analyses was assessed using permutation tests with the following procedure. For each ROI classification analysis the entire analysis was repeated 10,000 times with condition labels randomly assigned. Thus, we generated a null distribution of classification accuracies expected by chance and specific to each analysis. We assessed the significance of our ROI classification results by testing how often in the null distribution we obtained a mean decoding performance greater than or equal to the actual mean decoding performance found in our analysis. We first tested whether this was significant at the $p < 0.05$ threshold using a one-sided test (mean decoding performance in the null distribution must be greater than or equal to the actual mean decoding performance in less than 500 out of 10,000 tests). We then additionally applied a Bonferroni-correction to correct *p*-values for the number of ROIs tested. *P*-values following Bonferroni correction were limited to a maximum value of one (i.e. if Bonferroni correction caused a *p*-value to be greater than one, we set its value to one).

We additionally performed whole-brain searchlight analyses to investigate if any regions outside of our defined ROIs would be able to decode sex or weight. Searchlight analyses involved performing each classification analysis in 4-voxel radius spheres, where each sphere was centred around each voxel in the brain once, thus producing whole-brain maps of classification accuracies. To assess the significance of searchlight data we performed one-sided *t*-tests in SPM12 using a False Discovery Rate (FDR) correction to adjust for multiple comparisons.

2.7.1. MVPA analysis 1: classification of sex and weight

Our first set of analyses aimed to determine which brain regions encode the subcategories sex and weight, separately based on neural activity induced by viewing faces and bodies of these subcategories. For classification of sex, the classifier was trained and tested on male versus female stimuli, regardless of image size and weight. For classification of weight, the classifier was trained and tested on higher weight versus lower weight stimuli, regardless of image size and sex. Thus, we maximised the amount of data for the classifier to make its decision, and pooled across irrelevant subcategories. This meant also that the classifier was trained to be invariant to one of the features when distinguishing the other. We used 3 out of 4 runs as training data for the SVM classifier. We then used the trained classifier to predict the subcategories of the blocks in the independent test data from the 4th withheld run. A four-fold cross-validation procedure was used, such that each run was used once as the held out test dataset. Final decoding accuracy was determined by averaging over the four cross-validation iterations.

2.7.2. MVPA analysis 2: size-invariant classification of sex and weight

Our second set of analyses aimed to determine which brain regions encode the subcategories sex and weight in an image-size invariant manner. To this end, we trained the classifier to decode the subcategory from neural data evoked by one image size and then tested its ability to decode this subcategory from neural activity evoked by the other image size. As previously, 3 out of 4 runs were used as training data for the SVM classifier and the 4th withheld run was used as test data. However, in this case the training data only used neural activity data evoked by one image size, whereas the test data only used neural activity data evoked by subjects viewing the other image size. A four-fold cross-validation procedure was again used, such that each run was used once as the held out test dataset. This was repeated two times, once using the smaller image size data as training data and the larger image size data as test data, and vice-versa. The final decoding accuracy was determined by averaging over the four cross-validation iterations and both image size training and test set combinations.

2.7.3. MVPA analysis 3: classification of sex and weight across face and body stimuli

Our third set of analyses investigated which brain regions encode the subcategories sex and weight across face and body stimuli. To this end, we trained classifiers to decode each subcategory from neural data evoked by faces and tested them on neural activity evoked by bodies, and vice-versa. As face and body stimuli were shown in separate runs (4 runs each) we trained classifiers on all 4 runs of data evoked by one category (i.e. face or body stimuli) and tested them on all 4 runs of data evoked by the other category. The final decoding accuracy was determined by averaging over the two training and test set combinations.

2.8. Univariate analyses

We also performed univariate analyses to identify brain regions that are sensitive to the sex, weight, and image size of faces and bodies. In both separate and distributed ROIs we used *t*-tests to look for significant differences in the average BOLD response. *P*-values were Bonferroni-corrected for the number of ROIs tested. In addition, whole-brain analyses were conducted to look for additional brain regions with univariate activation differences, using a FDR correction to adjust for multiple comparisons.

2.9. Data and code availability statement

Data cannot be shared as participants were informed that their data would be stored confidentially, in accordance with the rules of the local ethics committee. Code is available on request.

3. Results

3.1. Behavioural results

3.1.1. Perceptual differences of face and body stimuli

Our selected face and body stimuli were perceptually different in terms of perceived sex and weight (Fig. 2). Participants gave significantly different ratings of sex to male and female faces and bodies (Fig. 2A; faces: $t_5 = 15.2$, $p = 2.3 \times 10^{-5}$; bodies: $t_5 = 20.0$, $p = 5.8 \times 10^{-6}$) and gave a significantly higher weight score for higher-weight stimuli than lower-weight stimuli (Fig. 2B; faces: $t_5 = 7.1$, $p = 8.8 \times 10^{-4}$; bodies: $t_5 = 9.3$, $p = 2.5 \times 10^{-4}$). There were also differences in the rating

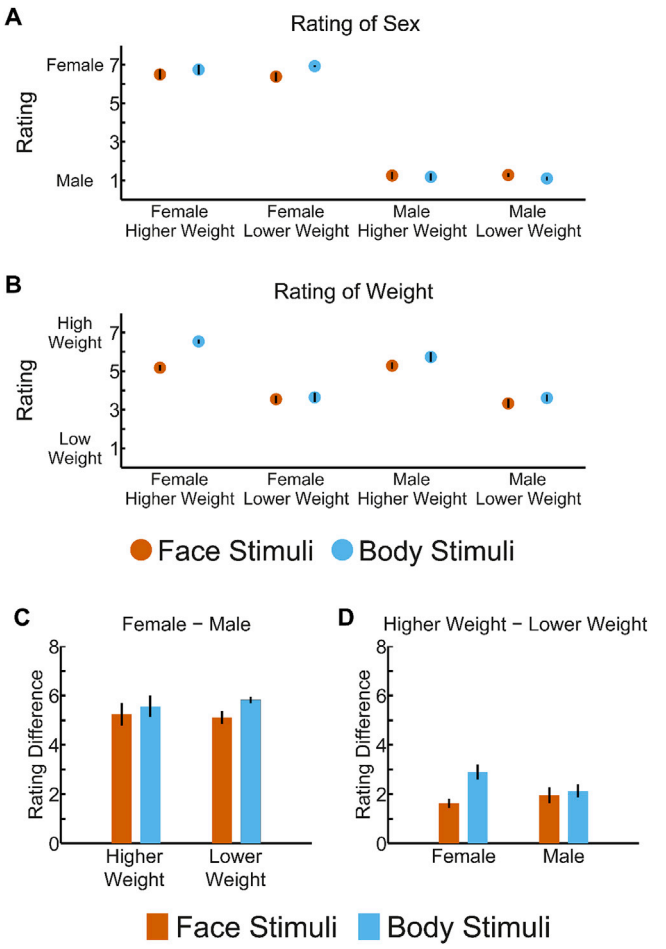

Fig. 2. Perceptual ratings of stimuli in an independent behavioural experiment. Participants rated how they perceived the biological sex (A) and weight (B) of face and body stimuli on 7-point scales. Circles indicate the mean rating scores and error bars show the standard error of the mean. (C) and (D) show the difference between the mean ratings of sex for the male and female stimuli (C), and between the mean ratings of weight for the higher and lower weight stimuli (D). Error bars show the standard error of the mean.

between face and body stimuli. The difference between male and female ratings was greater for body stimuli as compared to face stimuli ($t_5 = 5.3$, $p = 0.0031$; Fig. 2C) and the difference between higher and lower weight ratings was greater for body stimuli as compared to face stimuli ($t_5 = 2.8$, $p = 0.04$; Fig. 2D). Note that during the fMRI experiment, participants performed a simple task on the fixation cross that was unrelated to the stimuli.

3.1.2. Attention task during scanning

Participants showed high performance on the attention task during scanning (mean detection rate of 92% across all conditions). Detection performance showed no difference between face and body conditions ($t_{11} = 0.42$, $p = 0.68$), between male and female conditions ($t_{11} = -0.34$, $p = 0.74$), between higher and lower weight conditions ($t_{11} = 0.02$, $p = 0.98$) or between larger and smaller image size conditions ($t_{11} = -0.33$, $p = 0.74$). There was also no difference in reaction times between face and body conditions ($t_{11} = -0.36$, $p = 0.71$), between male and female conditions ($t_{11} = 0.06$, $p = 0.95$), between higher and lower weight conditions ($t_{11} = -0.03$, $p = 0.98$) or between larger and smaller image size conditions ($t_{11} = -0.21$, $p = 0.84$). Therefore, any observed decoding difference between experimental conditions (i.e. faces or bodies, sex, weight or image size) cannot be attributed to different levels of attention during scanning.

3.2. Classification of sex and weight

We first tested which brain regions show separable patterns of neural activity evoked by male and female stimuli or by higher- and lower-weight stimuli, regardless of image size. We used a leave-one-run-out cross-validation method, so that for each iteration we trained a linear support vector machine (SVM) classifier using data from three fMRI runs and tested the classifier with the data from the one remaining run. We performed these classification analyses in three types of brain regions; (1) separate face- and body-responsive regions, (2) distributed face- and body-responsive regions, and (3) in whole-brain searchlight analyses. The results are shown in Figs. 3 and 4.

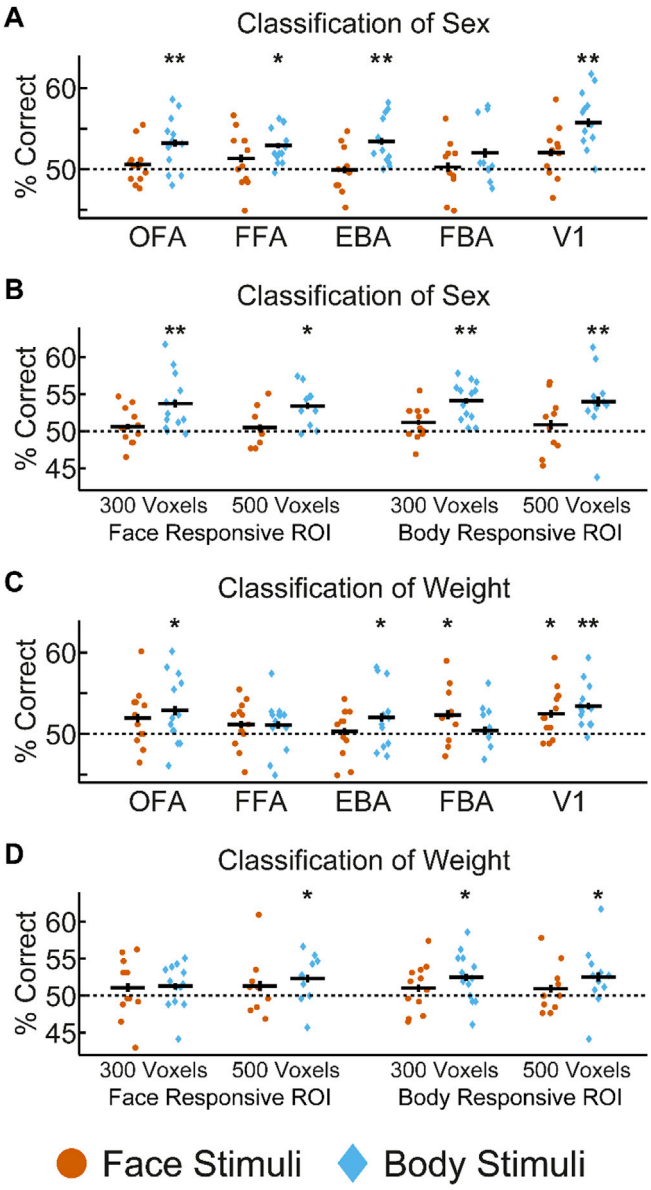

Fig. 3. Classification of sex and weight. Classification results for sex in separate ROIs (A) and distributed ROIs (B). Classification results for weight in separate ROIs (C) and distributed ROIs (D). Scatter plots show decoding accuracy for individual participants, horizontal black lines show group mean decoding accuracies and vertical error bars show the standard error of the mean. * indicates $p < 0.05$, ** indicates $p < 0.001$, Bonferroni-corrected for the number of ROIs (separate ROIs: N = 5; distributed ROIs: N = 4). Dotted lines indicate chance-level decoding performance, 50%.

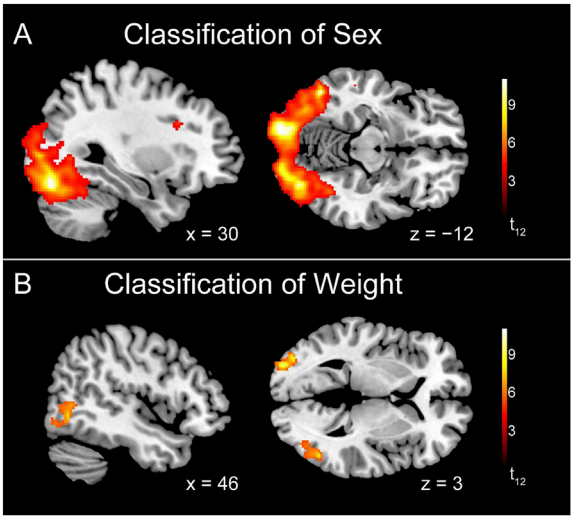

Fig. 4. Results of searchlight analyses for the classification of the sex (A) and weight (B) of bodies. (A) Regions able to classify the sex of bodies in the searchlight analysis overlapped with the mean coordinates of the rEBA, OFA, FFA, rFBA and V1. (B) Regions able to classify the weight of bodies in the searchlight analysis overlapped with the mean coordinates of the lOFA and rEBA.

3.2.1. Classification using separate ROIs

For body stimuli, classification of sex (Fig. 3A) was significantly above chance in EBA (53.4%, $p < 0.0005$), but not FBA (52.0%, $p = 0.078$). In addition, face-responsive areas OFA (53.2%, $p = 0.001$) and FFA (52.9%, $p = 0.0025$), as well as V1 (55.7%, $p < 0.0005$), also showed higher-than-chance classification of the sex of bodies. Classification of body weight (Fig. 3C) was significantly above chance in the EBA (52.0%, $p = 0.044$), but not the FBA (50.4%, $p = 1$). Again, we found higher-than chance level classification of body weight in OFA (52.9%, $p = 0.0015$) and V1 (53.4%, $p < 0.0005$). These results indicate that the encoding of body subcategories is not unique to body-responsive regions as both body- and face-responsive areas encode information about the sex and weight of bodies. Higher-than-chance decoding performance observed in V1 suggests that low-level visual features could be used to distinguish body subcategories.

For face stimuli, none of the separate ROIs was able to classify the sex of faces at a higher-than-chance level (Fig. 3A): OFA (50.6%, $p = 1$), FFA (51.3%, $p = 0.36$), EBA (49.9%, $p = 1$), FBA (50.3%, $p = 1$), V1 (52.1%, $p = 0.073$). Classification of the weight of faces was above chance in FBA (52.3%, $p = 0.046$) and V1 (52.5%, $p = 0.017$), but not in OFA (52.0%, $p = 0.072$), FFA (51.2%, $p = 0.48$), or EBA (50.3%, $p = 1$).

To examine whether our results were influenced by the way we defined ROIs (i.e. using a fixed number of voxels), we performed additional analyses using ROIs defined at a constant t-contrast threshold rather than a constant number of voxels. These additional analyses revealed the same results as reported above (see Supplementary Fig. S1), with the exception that the classification of the weight of faces was not significant in the FBA (50.9%, $p = 0.80$).

3.2.2. Classification using distributed ROIs

We further investigated classification of sex and weight in larger, distributed face- and body-responsive ROIs. These ROIs may show improved classification performance compared to smaller separated ROIs, due to increased information available for classification and known effects of ROI size on classification accuracy (Cox and Savoy, 2003; Gardumi et al., 2016). We defined ROIs of two different sizes: 300 voxels and 500 voxels (see Section 2.6 for details). For body stimuli, classification of sex (Fig. 3B) was higher than chance in both sizes of body-responsive ROI (300 voxels: 54.1%, $p < 0.0004$; 500 voxels: 54.0%, $p < 0.0004$), as well as both sizes of face-responsive ROI (300 voxels: 53.7%, $p < 0.0004$; 500 voxels: 53.4%, $p = 0.0024$). Classification of body weight (Fig. 3D) was also higher than chance in both sizes of body-responsive ROI (300 voxels: 52.5%, $p = 0.012$; 500 voxels: 52.5%, $p = 0.015$), and in the larger face-responsive ROI (500 voxels: 52.3%, $p = 0.03$).

For face stimuli, no distributed face or body ROIs were able to decode the sex or weight of faces, even using the large 500-voxel ROIs (which had the most information available for classification). Specifically, for the sex of faces the decoding performance was 50.5% ($p = 1$) in the larger face-responsive ROI and 50.9% ($p = 0.75$) in the larger body-responsive ROI. For the weight of faces the decoding performance was 51.3% ($p = 0.40$) in the larger face-responsive ROI and 50.9% ($p = 0.68$) in the larger body-responsive ROI. It is worth noting that when the distributed face-responsive ROI was defined at a constant t-contrast threshold rather than constant number of voxels, we were able to decode the weight of faces (52.2%, $p = 0.015$; see Supplementary Fig. S1D).

3.2.3. Classification using whole-brain searchlight analysis

To investigate if any other brain regions were sensitive to sex or weight information of faces and bodies, we performed the same classification analyses at the whole-brain level using a searchlight analysis. Fig. 4 shows the brain regions identified in the searchlight analyses with body stimuli. Higher-than-chance decoding of the sex of body stimuli was observed in much of occipitotemporal cortex (Fig. 4A) and consistent with the ROI classification analysis, the regions sensitive to the sex of bodies overlapped with the peak coordinates of body-responsive rEBA and rFBA, as well as face-responsive OFA and FFA and V1. Additional clusters sensitive to the sex of bodies were revealed in parietal (MNI coordinates: −2, −66, 56) and frontal regions (MNI coordinates: 40, 60, 6; 20, 32, 42; 28, 16, 28). These results show that many regions contain low-level visual, high-level visual or semantic information that allows a classifier to distinguish between neural activity evoked by bodies of different sexes.

Classification of the weight of body stimuli was significantly above chance in two regions that overlapped with the mean coordinates of lOFA and rEBA, as well as a small cluster in the right anterior temporal lobe (MNI coordinates: 40, −6, 32). For the face stimuli, classification of weight revealed a cluster in the left early visual cortex (MNI coordinates: −14, −94, −12). No regions were identified that could decode the sex of face stimuli.

3.3. Size-invariant classification of sex and weight

To investigate whether neural responses to sex and weight were invariant to image size, we tested whether patterns of neural activity evoked by one size of stimuli could be generalized to decode the sex or weight when participants view stimuli of a different size. Thus, we trained SVM classifiers to distinguish sex and weight using data obtained with only one of the two image sizes. We then tested the classifier using only the data obtained with the other image size. Significant decoding would suggest that the pattern of neural response to sex or weight was invariant to image size. Again, we performed the size invariant classification analysis using separate ROIs, distributed ROIs, and in searchlight analyses across the whole brain. The results are shown in Fig. 5.

For classification analysis using the separate ROIs, both body-responsive ROIs showed higher-than-chance performance when decoding the sex of body stimuli (EBA, 52.7%, $p = 0.002$; FBA, 52.8%, $p = 0.003$), but face-responsive ROIs OFA (50.9%, $p = 0.67$) and FFA (51.1%, $p = 0.47$), as well as V1 (50.2%, $p = 1$), did not. Paired-sample t-tests showed that decoding performance was significantly higher than V1 in EBA ($t_{12} = 3.2$, $p = 0.0077$) but not in FBA ($t_{10} = 1.6$, $p = 0.15$). Furthermore, classification results using ROIs defined at a constant t-contrast threshold (Supplementary Fig. S2A) showed significant decoding of the sex of bodies in EBA (53.0%, $p = 0.0005$), OFA (52.2%, $p = 0.016$) and V1 (52.3%, $p = 0.0070$) but not in FBA (51.1%, $p = 0.43$).

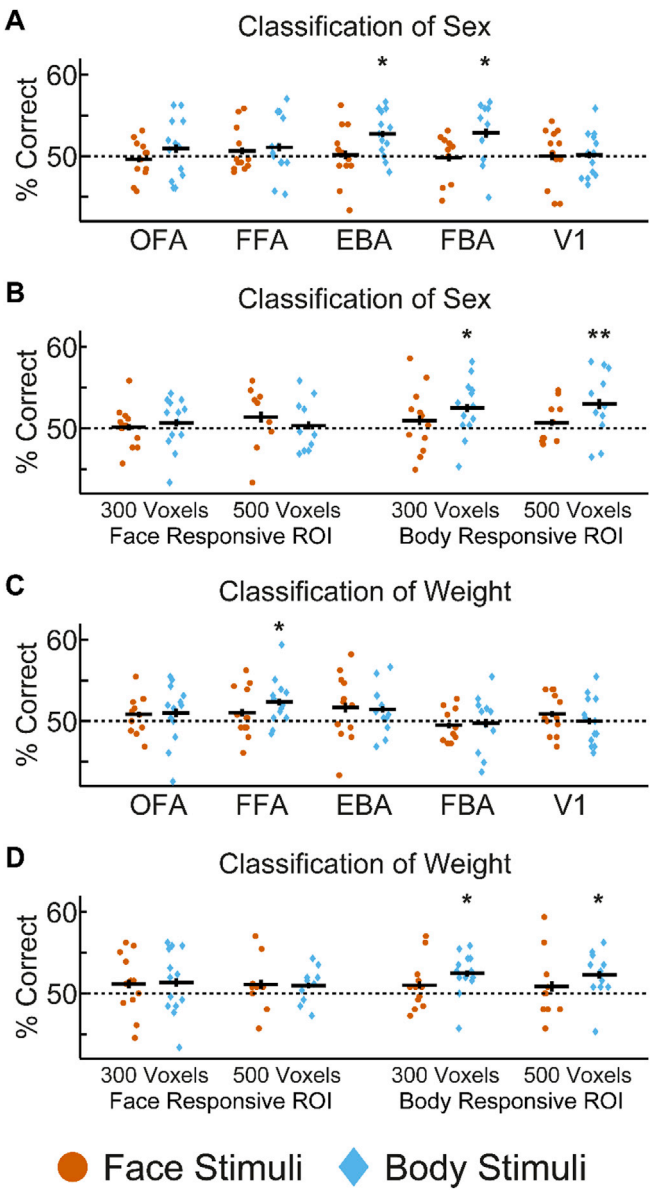

Fig. 5. Size-invariant classification of sex and weight. In this cross-classification analysis, a SVM classifier was trained on neural activity from subjects viewing stimuli of one image size and tested on its decoding performance on neural activity from the subjects viewing a different image size. (A) and (B) show classification results for sex in separate ROIs (A) and distributed ROIs (B). (C) and (D) show classification results for weight in separate ROIs (C) and distributed ROIs (D). Scatter plots show decoding accuracy for individual participants, horizontal black lines show group mean decoding accuracies and vertical error bars show the standard error of the mean.* indicates $p < 0.05$, ** indicates $p < 0.001$, Bonferroni-corrected for the number of ROIs (separate ROIs: N = 5; distributed ROIs: N = 4). Dotted lines indicate chance-level decoding performance, 50%.

These results suggest that decoding the sex of body stimuli across image size is most robust in the EBA.

Size-invariant classification of the weight of bodies was significant in FFA (52.3%, $p = 0.013$) but not in the body-related ROIs (EBA: 51.4%, $p = 0.20$; FBA: 49.8%, $p = 1$). A paired *t*-test showed that the classification of body weight in FFA was also significantly higher than in V1 ($t_{12} = 2.3$, $p = 0.042$). Classification results using ROIs defined at a constant t-contrast threshold (Supplementary Fig. S2C) showed significant decoding of body weight in OFA (52.3%, $p = 0.013$) but not in FFA (51.2%, $p = 0.40$). This difference may be due to an increase in information to the larger OFA ROI defined at a contrast t-contrast threshold, and an increase in noisy information hindering classification in the FFA ROI. No regions contained size-invariant information about the sex or weight of face stimuli.

For classification analyses using distributed ROIs, size-invariant classification of the sex of body stimuli was observed in both sizes of body-responsive ROI (300 voxels: 52.5%, $p = 0.0048$; 500 voxels: 53.0%, $p = 0.0008$) but not in face-responsive ROIs (300 voxels: 50.7%, $p = 0.85$; 500 voxels: 50.4%, $p = 1$). Similarly, size-invariant classification of the weight of bodies was found in both sizes of body-responsive ROI (300 voxels: 52.5%, $p = 0.0028$; 500 voxels: 52.3%, $p = 0.022$), but not in face-responsive ROIs (300 voxels: 51.4%, $p = 0.20$; 500 voxels: 51.0%, $p = 0.62$). Neither face- nor body-responsive ROIs could decode the sex or weight of face stimuli across image size (sex classification, face-responsive ROI, 500 voxels: 51.4%, $p = 0.32$; sex classification, body-responsive ROI, 500 voxels: 50.7%, $p = 0.94$; weight classification, face-responsive ROI, 500 voxels: 51.1%, $p = 0.57$; weight classification, body-responsive ROI, 500 voxels: 50.9%, $p = 0.73$).

The whole-brain searchlight analyses revealed no significant regions that could reliably decode the sex or weight of stimuli. This null-result, despite significant ROI decoding, might be due to the individual variance in ROI location that would weaken the group searchlight result in normalized space, combined with the overall weaker signal in the cross-decoding approach.

3.4. Classification of sex and weight across face and body stimuli

We investigated whether patterns of neural response to sex and weight would be able to generalize from patterns of activity evoked by faces to those evoked by bodies and vice-versa. To do this we trained SVM classifiers to distinguish sex and weight using neural activity data when participants viewed faces, and then tested the classifier on neural activity data when participants viewed bodies, and vice-versa. Significant decoding across face and body stimuli would suggest a semantic representation of the subcategory. We performed classification analyses with separate ROIs, distributed ROIs, and in searchlight analyses across the whole brain. The results are shown in Fig. 6.

We found significant decoding of weight across face and body stimuli in the EBA (51.5%, $p = 0.021$) and the 500 voxel body-responsive ROI (51.7%, $p = 0.016$), but not in any other ROIs (FBA: 50.6%, $p = 0.89$; OFA: 51.1%, $p = 0.14$; FFA: 49.8%, $p = 1$; V1: 49.4%, $p = 1$; face-responsive ROI, 500 voxels: 50.6%, $p = 0.84$). Paired-sample t-tests showed that decoding performance was significantly higher than V1 in EBA ($t_{12} = 3.3$, $p = 0.0068$) and the 500 voxel body-responsive ROI ($t_9 = 3.1$, $p = 0.012$). We additionally compared the classification performance of weight across face and body stimuli using ROIs defined at a constant t-contrast threshold (Supplementary Fig. S3C). We did not find significant decoding of weight in EBA (50.4%, $p = 1$) or the distributed body-responsive ROI (51.1%, $p = 0.063$) as defined using this method.

We were not able to decode sex across face and body stimuli in any of our separate or distributed ROIs (EBA: 50.0%, $p = 1$; FBA: 50.6%, $p = 0.9$; OFA: 50.4%, $p = 0.39$; FFA: 49.2%, $p = 1$; V1: 50.0%, $p = 1$; body-responsive ROI, 500 voxels: 50.1%, $p = 1$; face-responsive ROI, 500 voxels: 49.4%, $p = 1$). This non-significant decoding of sex across face and body stimuli is probably due to the fact that we were unable to identify any regions that could decode the sex of faces.

We additionally performed searchlight analyses to investigate if any other brain regions would be able to decode sex or weight across face and body stimuli. We did not identify any regions in these analyses.

3.5. Univariate analyses

To investigate whether the sex, weight, and image size of faces or bodies elicited different overall levels of neural activity, we conducted both ROI and whole-brain univariate analyses. The results are shown in Figs. 7–9.

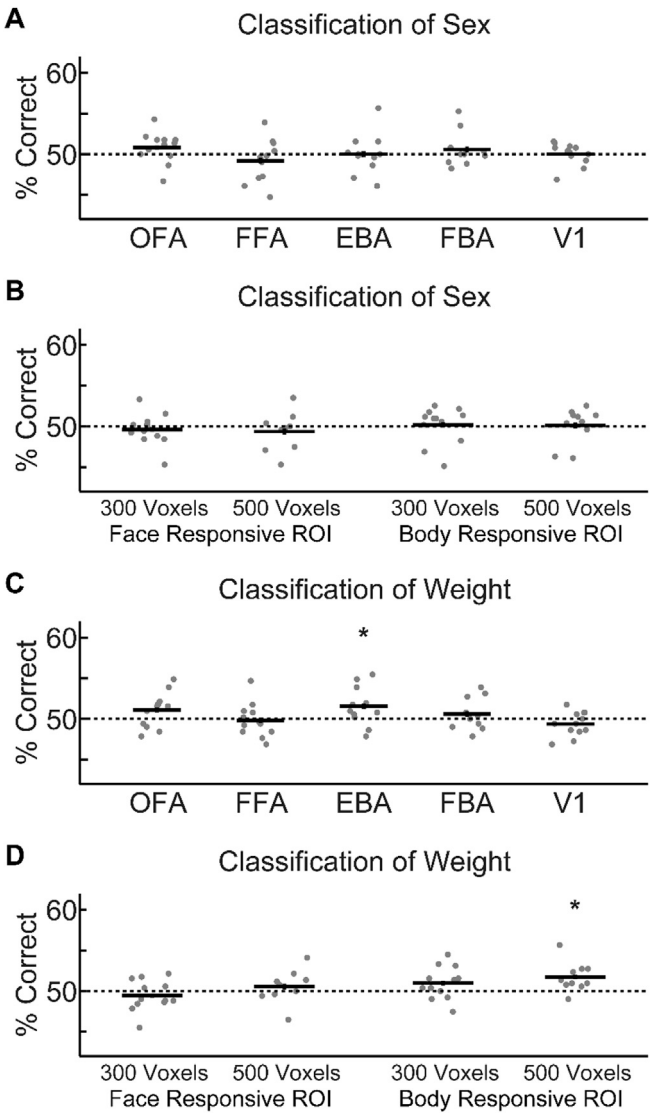

Fig. 6. Classification of sex and weight across face and body stimuli. In these cross-classification analyses, SVM classifiers were trained to distinguish sex and weight from neural activity when participants viewed faces and then subsequently tested on neural activity when participants viewed bodies, and vice-versa. (A) and (B) show classification results for sex in separate ROIs (A) and distributed ROIs (B). (C) and (D) show classification results for weight in separate ROIs (C) and distributed ROIs (D). Scatter plots show decoding accuracy for individual participants, horizontal black lines show group mean decoding accuracies and vertical error bars show the standard error of the mean.* indicates $p < 0.05$ Bonferroni-corrected for the number of ROIs (separate ROIs: N = 5; distributed ROIs: N = 4). Dotted lines indicate chance-level decoding performance, 50%.

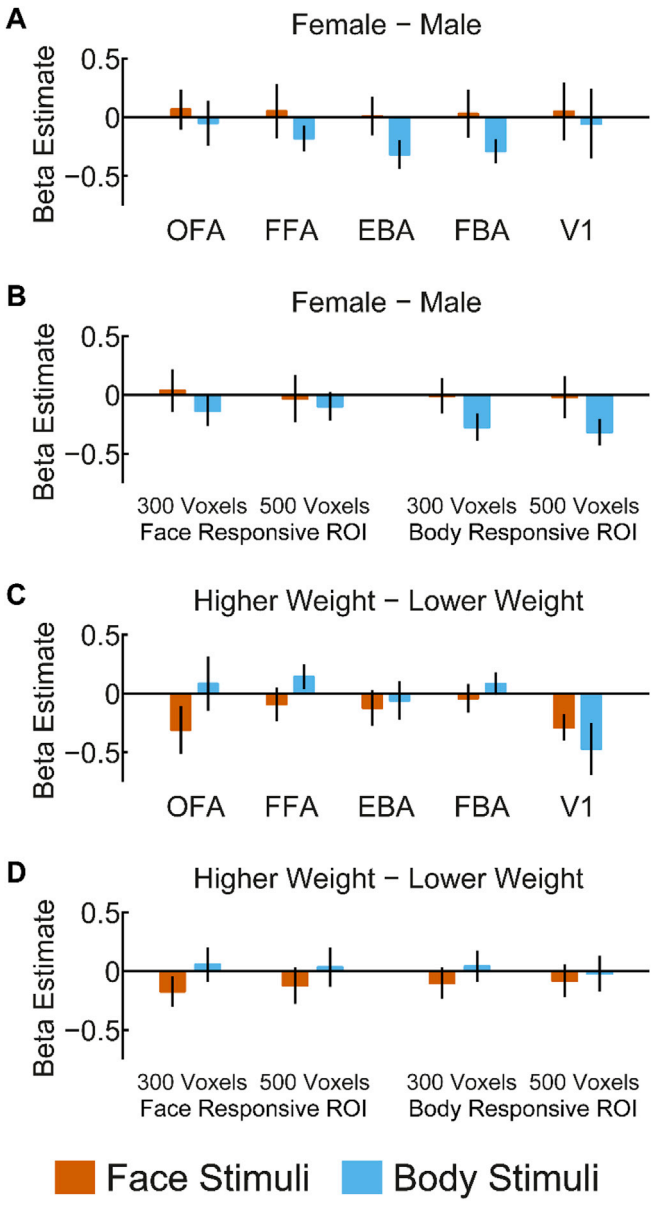

Fig. 7. Mean BOLD differences between the categories sex and weight. (A) and (B) illustrate mean differences between male and female stimuli for separate (A) and distributed (B) ROIs. Positive values indicate higher activation by female stimuli, negative values higher activation by male stimuli. (C) and (D) illustrate mean differences between higher and lower weight stimuli for separate (C) and distributed (D) ROIs. Positive values indicate higher activation by higher weight stimuli, negative values higher activation by lower weight stimuli. None of the differences were significant in any of the ROIs. Error bars show the standard error of the mean.

3.5.1. ROI analyses

Firstly, to investigate if there were any differences in mean response between faces or bodies of different sexes, we performed the contrast female stimuli minus male stimuli. None of the face- or body-responsive ROIs showed significant differences between male and female stimuli (Fig. 7A–B). There was a slight trend for higher activity to male bodies compared to female bodies in the body responsive regions, EBA ($t_{12} = -2.6$, $p = 0.12$), FBA ($t_{10} = -2.8$, $p = 0.098$), 300 voxel body-responsive ROI ($t_{12} = -2.4$, $p = 0.14$) and 500 voxel body-responsive ROI ($t_{10} = -2.9$, $p = 0.069$).

Secondly, we investigated if there were any differences in the mean response to faces or bodies of different weights, using the contrast higher weight stimuli minus lower weight stimuli. No significant differences between higher and lower weight stimuli were identified in any ROI (Fig. 7C–D). V1 showed a slight trend to higher activity to lower weight bodies compared to higher weight bodies ($t_{12} = -2.1$, $p = 0.27$) and to lower weight faces compared to higher weight faces ($t_{12} = -2.6$, $p = 0.12$).

Finally, we investigated if there were any differences in the mean response to faces or bodies of different image sizes (Fig. 8). For face stimuli, we found significantly higher BOLD activation to larger faces compared to smaller faces in FFA ($t_{11} = 8.2$, $p = 2.6 \times 10^{-5}$), FBA ($t_9 = 5.9$, $p = 0.0011$), both distributed face-responsive ROIs (300 voxels: $t_{11} = 3.8$, $p = 0.012$; 500 voxels: $t_8 = 3.7$, $p = 0.024$) and both distributed body-responsive ROIs (300 voxels: $t_{11} = 3.8$, $p = 0.013$; 500 voxels:

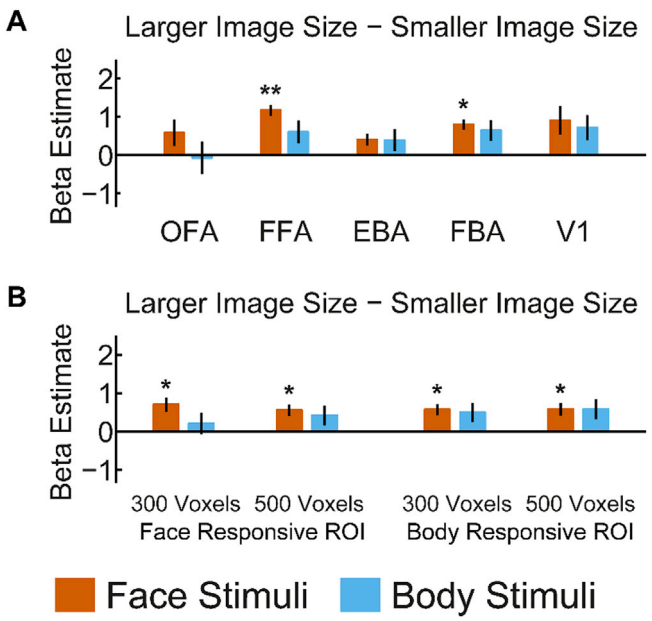

Fig. 8. Mean BOLD differences between larger and smaller images. (A) and (B) illustrate mean differences between larger and smaller image size stimuli for separate (A) and distributed (B) ROIs. Positive values indicate higher activation by larger size stimuli, negative values higher activation by smaller size stimuli. * indicates $p < 0.05$, ** indicates $p < 0.001$, Bonferroni-corrected for the number of ROIs (separate ROIs: N = 5; distributed ROIs: N = 4). Error bars show the standard error of the mean.

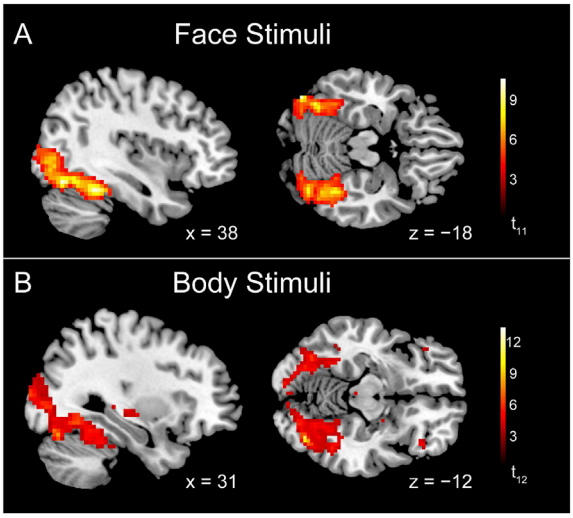

Fig. 9. Whole-brain results showing univariate activation differences between larger and smaller image size stimuli. (A) Brain regions showing higher activation to larger faces compared to smaller faces. These regions overlapped with the mean coordinates of the OFA, FFA and FBA bilaterally. (B) Brain regions showing higher activation to larger bodies compared to smaller bodies. These regions overlapped with the mean coordinates of the FBA, rEBA and rFFA.

$t_9 = 3.5$, $p = 0.027$). Additionally, a slight trend to higher activity to larger compared to smaller faces was seen in EBA ($t_{11} = 2.7$, $p = 0.10$) and V1 ($t_{11} = 2.4$, $p = 0.17$). For body stimuli, none of ROIs showed significant differences between larger and smaller bodies. A trend for higher activity to larger than smaller bodies was seen in FBA ($t_{10} = 2.4$, $p = 0.19$), V1 ($t_{12} = 2.2$, $p = 0.25$) and the 500-voxel body-responsive ROI ($t_{10} = 2.2$, $p = 0.20$).

3.5.2. Whole-brain analyses

To investigate if any other brain regions were sensitive to the sex, weight, and image size of faces and bodies, we performed the same three contrasts in whole-brain analyses. We found that the occipitotemporal cortex showed higher brain activity for larger stimuli than smaller stimuli, for both faces (Fig. 9A) and bodies (Fig. 9B). For face stimuli the significant area overlapped with the mean coordinates of the OFA, FFA and FBA bilaterally. For body stimuli the significant area overlapped with the mean coordinates of the FBA, rEBA and rFFA. As for the sex and weight of stimuli, no brain regions showed significant differences between male and female stimuli or between higher- and lower-weight stimuli.

4. Discussion

In this study, we investigated how face- and body-responsive brain regions encode information about the subcategories sex and weight. We show, for the first time to our knowledge, that subcategorical information about bodies is encoded in both body- and face-responsive areas in the brain, and that this information is encoded in a size-invariant manner, more so for the body-than face-related brain network. Furthermore, we find evidence that the FBA responds to the weight of faces, and that weight is encoded in an abstract manner in the EBA and distributed body-responsive network. These results indicate that subcategories shared by faces and bodies (e.g. sex) are encoded by different patterns of neural responses in the brain network related to person perception.

4.1. Neural coding of body subcategories

We found that both face-responsive and body-responsive regions encoded information about the sex and weight of bodies. For the sex of bodies, higher-than-chance decoding was observed not only in all body-responsive ROIs but also in all face-responsive ROIs. For the weight of bodies, both the distributed body-responsive areas and the largest-size distributed face-responsive area showed higher-than-chance classification of body weight. For the separate ROI analyses, higher-than-chance decoding of body weight was observed in the EBA as well as the face-responsive ROIs OFA and FFA. In contrast to the differential multivariate activity patterns, no ROI showed stronger net activity for one subcategory than another.

Our searchlight results showed that brain regions responding to the weight of bodies overlapped with those responding to the sex of bodies, though the latter is more widespread. This result shows similarity to a recent behavioural finding that judgements of the sex of bodies are independent of judgements of the weight of bodies but not vice-versa (Johnstone and Downing, 2017). This coincidence suggests that the above behavioural difference might be related to how the brain regions processing these two subcategories overlap. The encoding of body subcategory information in face-responsive areas indicates that these face areas are not exclusively involved in face processing (cf. Kanwisher and Yovel, 2006). These results are in line with previous findings showing that categorical (e.g. faces and bodies) and subcategorical information is distributed across occipitotemporal cortex, rather than selective to specific sub-regions of cortex (Haxby et al., 2001; Huth et al., 2012; Op de Beeck et al., 2010).

Although both face- and body-responsive brain areas encoded the sex and weight of bodies, size-invariant body information was more prominent and consistent in the body-responsive areas. When we used data obtained from one image size to train the classifier and then tested it with the data obtained from a different image size, we were able to decode the sex of bodies from EBA, FBA and both distributed body-responsive regions, but not from any face-responsive region. Similarly, the two distributed body-responsive ROIs, but not face-responsive ROIs, showed higher-than-chance decoding of body weight in a size-invariant manner. Although the FFA showed size-invariant decoding of the weight of bodies, in total less face-responsive ROIs showed size-invariant decoding

of body subcategories than body-responsive ROIs.

In our first classification analysis we found that V1 was able to decode the sex of bodies and the weight of bodies and faces, suggesting there may be some low-level visual information that could be used for classification in these analyses (where both image-sizes were present in both the training and test set). In contrast, size-invariant decoding of sex and weight was not possible in our V1 ROI.

4.2. Neural coding of face subcategories

We were able to decode the weight of faces from the FBA, and classify weight across face and body stimuli in the EBA and distributed body-responsive network. The decoding of weight across face and body stimuli in the EBA and distributed body-responsive network suggests that these regions may contain a semantic encoding of weight (i.e. not dependent upon low-level visual features). Although the sex of faces can be clearly differentiated in perception (as shown by perceptual ratings), we were unable to decode the sex of faces based on the pattern of neural activity recorded with fMRI. Previous studies have found mixed results about whether the sex of faces can be decoded based on the patterns of neural activity in face-responsive brain areas. Some studies were not able to decode the sex of faces from the occipitotemporal cortex (Kriegeskorte et al., 2007) and the FFA (Kanwisher, 2017), whereas others showed that the sex of faces can be successfully decoded from the FFA or extended face network (Contreras et al., 2013; Kaul et al., 2011). One factor that may cause such discrepant results is the stimuli. Our face stimuli were carefully controlled and lacked external face features (such as hair or make-up) whereas Contreras et al. (2013) and Kaul et al. (2011) used photographs of faces varying in hairstyle, make-up, and beards. These external cues are often diagnostic for the sex of faces and facilitate sex categorization (Rossion, 2002). Thus, despite clear perceptual differences shown in the rating task, the sex differences in our face stimuli might be insufficient to elicit distinct patterns of brain activity that can be detected and decoded using MVPA. Note that a lack of ability to decode face subcategories does not necessarily reflect a lack of information about the subcategories in these regions, it may simply be beyond the resolution that fMRI MVPA can detect. For instance (Dubois et al., 2015), compared identity decoding from fMRI recordings with that from single cell recordings in the anterior temporal cortex of macaques: they found that identity information decodable from single cells was not decodable from the fMRI data.

Another factor that might have hindered our ability to decode the sex of faces was the task. Our attention task (i.e. detecting a red dot) did not involve any effortful processing of faces, whereas both Contreras et al. (2013) and Kaul et al. (2011) employed tasks encouraging effortful face processing (i.e. judgments about the gender or attractiveness of faces). Given that task-related attention has been shown to affect the neural representation of faces and semantic categories in natural vision (Çukur et al., 2013; Dobs et al., 2018; Kaiser et al., 2016), the automatic processing of faces in our study may lead to weak responses to the sex of faces, thereby reducing the possibility of higher-than-chance decoding. In addition, while sex and weight are two of the most salient dimensions that differentiate the shape of bodies (Hill et al., 2016), they are probably not the most important ones for faces (e.g. in comparison with the identity or emotional expression of faces (Burton et al., 1999; O'Toole et al., 2011)). In line with this, we found greater differences in perceptual ratings between male and female stimuli and between higher and lower weight stimuli for body stimuli than for face stimuli. We note however that lower rating differences cannot be the only reason for the differences we see in decoding performance as rating differences of the sex of faces (for which we do not find significant decoding in any region) are considerably larger than for the weight of bodies and faces (for which we find significant decoding in a number of brain regions). Lastly, our relatively small sample size ($N = 12$ for face stimuli) might also have impacted our ability to decode the sex of faces. A larger sample size might be able to improve the decoding of the sex of faces.

4.3. Neural coding of subcategories shared by faces and bodies

Previous studies investigating neural processing of face and body information have suggested that such information is supported by largely separated neural networks in occipitotemporal cortex (Peelen and Downing, 2007; Pitcher et al., 2009). This segregated neural processing of faces and bodies has been demonstrated with both human and nonhuman primates (Premereur et al., 2016). For example, neuroimaging studies in macaques and humans have shown that responses to whole-agents (i.e. where the image contains both the face and body) can be best modelled by a linear combination of responses to faces and bodies alone, suggesting that there are separate neural populations responsive to face and body information respectively (Fisher and Freiwald, 2015; Kaiser et al., 2014). In contrast to this parallel processing hypothesis, we found that the EBA and distributed body-responsive network could classify weight across face and body stimuli, suggesting that these regions encode some shared subcategorical information from faces and bodies in an abstract manner. This abstract coding may be semantic or based on high-level visual features (for example concavity vs convexity). Although sex can also be perceived from both faces and bodies, we were unable to identify brain regions showing stimuli-independent encoding of sex. In behaviour, psychological adaptation studies have demonstrated that perception of gender (Ghuman et al., 2010; Palumbo et al., 2014) can adapt across face and body stimuli, suggesting an overlapping representation.

Recent studies have suggested that the anterior temporal lobes may contain an integrated representation of body and face information. For example, the right ATFA showed an equivalent level of brain activity to faces and headless bodies and exhibited significant correlation between face- and body-elicited neural responses (Harry et al., 2016). Similarly, the anterior temporal face patch (ATFP) in macaques has been shown to respond more strongly to whole-agents than the addition of the responses to the face and body shown separately, suggesting an integration of face and body information in this region (Fisher and Freiwald, 2015). Such integration may help differentiate individuals' identity, as this area has been linked to identity representation in humans (Anzellotti et al., 2014; Guntupalli et al., 2016; Kriegeskorte et al., 2007). However, our results (from searchlight analyses across image size or across face and body stimuli) suggest that this identity-sensitive area does not automatically encode the sex or weight of a person in an abstract manner.

4.4. Effect of low-level stimulus features

We found consistently enhanced neural responses to faces and bodies when the size of images increased. For both faces and bodies, significantly higher brain activity for larger than smaller stimuli was observed in distributed areas across occipitotemporal cortex, covering face- and body-responsive ROIs, respectively. For face stimuli, the active area overlapped with the mean coordinates of face-responsive OFA and FFA, as well as body-responsive FBA. For body stimuli the overlap was with the mean coordinates of body-responsive FBA and rEBA, as well as face-responsive rFFA. Neural responses to faces appeared to be more sensitive to the stimulus size than those to bodies (see Figs. 8 and 9). The influence of stimulus size on neural response of high-level visual areas has been shown for faces (Yue et al., 2011) and general everyday objects (Konkle and Oliva, 2012), but not, to our knowledge, for human bodies. These results demonstrate that neural responses to faces and bodies in high-level visual areas are modulated by low-level stimulus properties, rather than being size-invariant (cf. Andrews and Ewbank, 2004; Sawamura et al., 2005).

5. Conclusion

Our study provides the first evidence, to our knowledge, that the sex and weight of human bodies can be decoded from neural activity in the person-related brain network using MVPA. By demonstrating size-

invariant decoding of body subcategories in both the body- and face-responsive brain network, as well as cross-classification of weight across face and body stimuli, we show that the neural responses to these subcategories are largely driven by high-level visual or semantic features rather than by merely low-level visual features. The present study also offers an alternative approach to repetition suppression for investigating neural responses to subcategorical body information using fMRI. Methods like repetition suppression may be biased by top-down effects (Summerfield et al., 2008). Together, these findings not only offer new insights into how the brain encodes person-related visual information like faces and bodies, but also shed light on how shared subcategories from visually distinctive object categories may be encoded in the human brain.

Acknowledgements

This research was supported by the Max Planck Society, Germany.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuroimage.2019.116085.

References

Aguirre, G.K., 2007. Continuous carry-over designs for fMRI. Neuroimage 35, 1480–1494. https://doi.org/10.1016/j.neuroimage.2007.02.005.

Andrews, T.J., Ewbank, M.P., 2004. Distinct representations for facial identity and changeable aspects of faces in the human temporal lobe. Neuroimage 23, 905–913. https://doi.org/10.1016/j.neuroimage.2004.07.060.

Anzellotti, S., Fairhall, S.L., Caramazza, A., 2014. Decoding representations of face identity that are tolerant to rotation. Cerebr. Cortex 24, 1988–1995. https://doi.org/10.1093/cercor/bht046.

Axelrod, V., Yovel, G., 2015. Successful decoding of famous faces in the fusiform face area. PLoS One 10, 1–20. https://doi.org/10.1371/journal.pone.0117126.

Blanz, V., Vetter, T., 1999. A morphable model for the synthesis of 3D faces. SIGGRAPH'99 Conf. Proc. 187–194. https://doi.org/10.1145/311535.311556.

Bracci, S., Op de Beeck, H., 2016. Dissociations and associations between shape and category representations in the two visual pathways. J. Neurosci. 36, 432–444. https://doi.org/10.1523/JNEUROSCI.2314-15.2016.

Brainard, D.H., 1997. The Psychophysics toolbox. Spat. Vis. 10, 433–436. https://doi.org/10.1163/156856897X00357.

Brooks, J.L., 2012. Counterbalancing for serial order carryover effects in experimental condition orders. Psychol. Methods 17, 600–614. https://doi.org/10.1037/a0029310.

Burton, A.M., Wilson, S., Cowan, M., Bruce, V., 1999. Face recognition in poor quality video. Psychol. Sci. 10, 243–248. https://doi.org/10.1111/1467-9280.00144.

Cohen, L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, M., Michel, F., 2000. The visual word form area: spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. Brain 123, 291–307.

Contreras, J.M., Banaji, M.R., Mitchell, J.P., 2013. Multivoxel patterns in fusiform face area differentiate faces by sex and race. PLoS One 8, 1–6. https://doi.org/10.1371/journal.pone.0069684.

Cox, D.D., Savoy, R.L., 2003. Functional magnetic resonance imaging (fMRI) "brain reading": detecting and classifying distributed patterns of fMRI activity in human visual cortex. Neuroimage 19, 261–270. https://doi.org/10.1016/S1053-8119(03)00049-1.

Çukur, T., Nishimoto, S., Huth, A.G., Gallant, J.L., 2013. Attention during natural vision warps semantic representation across the human brain. Nat. Neurosci. 16, 763–770. https://doi.org/10.1038/nn.3381.

Dobs, K., Schultz, J., Bülthoff, I., Gardner, J.L., 2018. Task-dependent enhancement of facial expression and identity representations in human cortex. Neuroimage 172, 689–702. https://doi.org/10.1016/j.neuroimage.2018.02.013.

Downing, P.E., Jiang, Y., Shuman, M., Kanwisher, N., 2001. A cortical area selective for visual processing of the human body. Science 293, 2470–2473. https://doi.org/10.1126/science.1063414.

Dubois, X.J., de Berker, A.O., Tsao, D.Y., 2015. Single-unit recordings in the macaque face patch system reveal limitations of fMRI MVPA. J. Neurosci. 35, 2791–2802. https://doi.org/10.1523/JNEUROSCI.4037-14.2015.

Epstein, R., Kanwisher, N., 1998. A cortical representation of the local visual environment. Nature 392, 598–601. https://doi.org/10.1038/33402.

Ewbank, M.P., Lawson, R.P., Henson, R.N., Rowe, J.B., Passamonti, L., Calder, A.J., 2011. Changes in "Top-Down" connectivity underlie repetition suppression in the ventral visual pathway. J. Neurosci. 31, 5635–5642. https://doi.org/10.1523/JNEUROSCI.5013-10.2011.

Fisher, C., Freiwald, W.A., 2015. Whole-agent selectivity within the macaque face-processing system. Proc. Natl. Acad. Sci. 112, 14717–14722. https://doi.org/10.1073/pnas.1512378112.

Freeman, J.B., Rule, N.O., Adams, R.B., Ambady, N., 2010. The neural basis of categorical face perception: graded representations of face gender in fusiform and orbitofrontal cortices. Cerebr. Cortex 20, 1314–1322. https://doi.org/10.1093/cercor/bhp195.

Gardumi, A., Ivanov, D., Hausfeld, L., Valente, G., Formisano, E., Uludağ, K., 2016. The effect of spatial resolution on decoding accuracy in fMRI multivariate pattern analysis. Neuroimage 132, 32–42. https://doi.org/10.1016/j.neuroimage.2016.02.033.

Gauthier, I., Tarr, M.J., Moylan, J., Skudlarski, P., Gore, J.C., Anderson, A.W., 2000. The fusiform "face area" is part of a network that processes faces at the individual level. J. Cogn. Neurosci. 12, 495–504. https://doi.org/10.1162/089892900562165.

Ghuman, A.S., McDaniel, J.R., Martin, A., 2010. Face adaptation without a face. Curr. Biol. 20, 32–36. https://doi.org/10.1016/j.cub.2009.10.077.

Grill-Spector, K., Weiner, K.S., 2014. The functional architecture of the ventral temporal cortex and its role in categorization. Nat. Rev. Neurosci. 15, 536–548. https://doi.org/10.1038/nrn3747.

Gross, C.G., Roch-Miranda, C.E., Bender, D.B., 1972. Visual properties of neurons in inferotemporal of the macaque. J. Neurophysiol. 35, 96–111.

Guntupalli, J.S., Wheeler, K.G., Gobbini, M.I., 2016. Disentangling the representation of identity from head view along the human face processing pathway. Cerebr. Cortex 27, 46–53. https://doi.org/10.1093/cercor/bhw344.

Hahn, C.A., O'Toole, A.J., 2017. Recognizing approaching walkers: neural decoding of person familiarity in cortical areas responsive to faces, bodies, and biological motion. Neuroimage 146, 859–868. https://doi.org/10.1016/j.neuroimage.2016.10.042.

Harry, B., Umla-Runge, K., Lawrence, A., Graham, K., Downing, P., 2016. Evidence for integrated visual face and body representations in the anterior temporal lobes. J. Cogn. Neurosci. 28, 1178–1193. https://doi.org/10.1162/jocn_a_00966.

Haxby, J.V., Gobbini, M.I., Furey, M.L., Ishai, A., Schouten, J.L., Pietrini, P., 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science 293, 2425–2430. https://doi.org/10.1126/science.1063736.

Hebart, M.N., Görgen, K., Haynes, J.-D., 2015. The Decoding Toolbox (TDT): a versatile software package for multivariate analyses of functional imaging data. Front. Neuroinf. 8, 1–18. https://doi.org/10.3389/fninf.2014.00088.

Hill, M.Q., Streuber, S., Hahn, C.A., Black, M.J., O'Toole, A.J., 2016. Creating body shapes from verbal descriptions by linking similarity spaces. Psychol. Sci. 27, 1486–1497. https://doi.org/10.1177/0956797616663878.

Hinds, O., Polimeni, J.R., Rajendran, N., Balasubramanian, M., Amunts, K., Zilles, K., Schwartz, E.L., Fischl, B., Triantafyllou, C., 2009. Locating the functional and anatomical boundaries of human primary visual cortex. Neuroimage 46, 915–922. https://doi.org/10.1016/j.neuroimage.2009.03.036.

Hummel, D., Rudolf, A.K., Brandi, M.L., Untch, K.H., Grabhorn, R., Hampel, H., Mohr, H.M., 2013. Neural adaptation to thin and fat bodies in the fusiform body area and middle occipital gyrus: an fMRI adaptation study. Hum. Brain Mapp. 34, 3233–3246. https://doi.org/10.1002/hbm.22135.

Huth, A.G., Nishimoto, S., Vu, A.T., Gallant, J.L., 2012. A continuous semantic space describes the representation of thousands of object and action categories across the human brain. Neuron 76, 1210–1224. https://doi.org/10.1016/j.neuron.2012.10.014.

Jeong, S.K., Xu, Y., 2016. Behaviorally relevant abstract object identity representation in the human parietal cortex. J. Neurosci. 36, 1607–1619. https://doi.org/10.1523/JNEUROSCI.1016-15.2016.

Johnstone, L.T., Downing, P.E., 2017. Dissecting the visual perception of body shape with the Garner selective attention paradigm. Vis. Cogn. 25, 507–523. https://doi.org/10.1080/13506285.2017.1334733.

Kaiser, D., Oosterhof, N.N., Peelen, M.V., 2016. The neural dynamics of attentional selection in natural scenes. J. Neurosci. 36, 10522–10528. https://doi.org/10.1523/JNEUROSCI.1385-16.2016.

Kaiser, D., Strnad, L., Seidl, K.N., Kastner, S., Peelen, M.V., 2014. Whole person-evoked fMRI activity patterns in human fusiform gyrus are accurately modeled by a linear combination of face- and body-evoked activity patterns. J. Neurophysiol. 111, 82–90. https://doi.org/10.1152/jn.00371.2013.

Kanwisher, N., 2017. The quest for the FFA and where it led. J. Neurosci. 37, 1056–1061. https://doi.org/10.1523/JNEUROSCI.1706-16.2016.

Kanwisher, N., McDermott, J., Chun, M.M., 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. J. Neurosci. 17, 4302–4311.

Kanwisher, N., Yovel, G., 2006. The fusiform face area: a cortical region specialized for the perception of faces. Philos. Trans. R. Soc. Biol. Sci. 361, 2109–2128. https://doi.org/10.1098/rstb.2006.1934.

Kaul, C., Rees, G., Ishai, A., 2011. The gender of face stimuli is represented in multiple regions in the human brain. Front. Hum. Neurosci. 4, 1–12. https://doi.org/10.3389/fnhum.2010.00238.

Kleiner, M., Brainard, D., Pelli, D., 2007. What's new in Psychtoolbox-3?. In: Perception 36 ECVP Abstract Supplement.

Konkle, T., Oliva, A., 2012. A real-world size organization of object responses in occipitotemporal cortex. Neuron 74, 1114–1124. https://doi.org/10.1016/j.neuron.2012.04.036.

Kriegeskorte, N., Formisano, E., Sorger, B., Goebel, R., 2007. Individual faces elicit distinct response patterns in human anterior temporal cortex. Proc. Natl. Acad. Sci. 104, 20600–20605. https://doi.org/10.1073/pnas.0705654104.

Kriegeskorte, N., Mur, M., Ruff, D.A., Kiani, R., Bodurka, J., Esteky, H., Tanaka, K., Bandettini, P.A., 2008. Matching categorical object representations in inferior temporal cortex of man and monkey. Neuron 60, 1126–1141. https://doi.org/10.1016/j.neuron.2008.10.043.

Logothetis, N.K., Sheinberg, D.L., 1996. Visual object recognition. Annu. Rev. Neurosci. 19, 577–621.

Loper, M., Mahmood, N., Romero, J., Pons-moll, G., Black, M.J., 2015. SMPL : a skinned multi-person linear model. ACM Trans. Graph. (Proc. SIGGRAPH Asia) 34 (248), 1–248, 16. https://doi.org/10.1145/2816795.2818013.

Malach, R., Reppas, J.B., Benson, R.R., Kwong, K.K., Jiang, H., Kennedy, W.A., Ledden, P.J., Brady, T.J., Rosen, B.R., Tootell, R.B.H., 1995. Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. Proc. Natl. Acad. Sci. U.S.A. 92, 8135–8139. https://doi.org/10.1073/pnas.92.18.8135.

O'Toole, A.J., Phillips, P.J., Weimer, S., Roark, D.A., Ayyad, J., Barwick, R., Dunlop, J., 2011. Recognizing people from dynamic and static faces and bodies: dissecting identity with a fusion approach. Vis. Res. 51, 74–83. https://doi.org/10.1016/j.visres.2010.09.035.

Op de Beeck, H.P., Brants, M., Baeck, A., Wagemans, J., 2010. Distributed subordinate specificity for bodies, faces, and buildings in human ventral visual cortex. Neuroimage 49, 3414–3425. https://doi.org/10.1016/j.neuroimage.2009.11.022.

Palumbo, R., D'Ascenzo, S., Tommasi, L., 2014. Cross-category adaptation: exposure to faces produces gender aftereffects in body perception. Psychol. Res. 79, 380–388. https://doi.org/10.1007/s00426-014-0576-2.

Peelen, M.V., Downing, P.E., 2007. The neural basis of visual body perception. Nat. Rev. Neurosci. 8, 636–648. https://doi.org/10.1038/nrn2195.

Peelen, M.V., Downing, P.E., 2005. Selectivity for the human body in the fusiform gyrus. J. Neurophysiol. 93, 603–608. https://doi.org/10.1152/jn.00513.2004.

Pitcher, D., Charles, L., Devlin, J.T., Walsh, V., Duchaine, B., 2009. Triple dissociation of faces, bodies, and objects in extrastriate cortex. Curr. Biol. 19, 319–324. https://doi.org/10.1016/j.cub.2009.01.007.

Premereur, E., Taubert, J., Janssen, P., Vogels, R., Vanduffel, W., 2016. Effective connectivity reveals largely independent parallel networks of face and body patches. Curr. Biol. 26, 3269–3279. https://doi.org/10.1016/j.cub.2016.09.059.

Ratner, K.G., Kaul, C., Van Bavel, J.J., 2013. Is race erased? decoding race from patterns of neural activity when skin color is not diagnostic of group boundaries. Soc. Cogn. Affect. Neurosci. 8, 750–755. https://doi.org/10.1093/scan/nss063.

Robinette, K.M., Blackwell, S., Daanen, H., Boehmer, M., Fleming, S., Brill, T., Hoeferlin, D., Burnsides, D., 2002. Civilian American and European Surface Anthropometry Resource (CAESAR), Final Report. US Air Force Research Laboratory. Tech. Rep. AFRL-HE- WP-TR-2002-0169.

Rossion, B., 2002. Is sex categorization from faces really parallel to face recognition? Vis. Cogn. 9, 1003–1020. https://doi.org/10.1080/13506280143000485.

Sawamura, H., Georgieva, S., Vogels, R., Vanduffel, W., Orban, G.A., 2005. Using functional magnetic resonance imaging to assess adaptation and size invariance of shape processing by humans and monkeys. J. Neurosci. 25, 4294–4306. https://doi.org/10.1523/JNEUROSCI.0377-05.2005.

Schwarzlose, R.F., Baker, C.I., Kanwisher, N., 2005. Separate face and body selectivity on the fusiform gyrus. J. Neurosci. 25, 11055–11059. https://doi.org/10.1523/JNEUROSCI.2621-05.2005.

Sturman, D., Stephen, I.D., Mond, J., Stevenson, R.J., Brooks, K.R., 2017. Independent Aftereffects of Fat and Muscle: implications for neural encoding, body space representation, and body image disturbance. Sci. Rep. 7, 1–8. https://doi.org/10.1038/srep40392.

Summerfield, C., Trittschuh, E.H., Monti, J.M., Mesulam, M.M., Egner, T., 2008. Neural repetition suppression reflects fulfilled perceptual expectations. Nat. Neurosci. 11, 1004–1006. https://doi.org/10.1038/nn.2163.

Tanaka, K., 1996. Inferotemporal cortex and object vision. Annu. Rev. Neurosci. 19, 109–139. https://doi.org/10.1146/annurev.neuro.19.1.109.

Troje, N.F., Bülthoff, H.H., 1996. Face recognition under varying poses: the role of texture and shape. Vis. Res. 36, 1761–1771. https://doi.org/10.1016/0042-6989(95)00230-8.

Yue, X., Cassidy, B.S., Devaney, K.J., Holt, D.J., Tootell, R.B.H., 2011. Lower-level stimulus features strongly influence responses in the fusiform face area. Cerebr. Cortex 21, 35–47. https://doi.org/10.1093/cercor/bhq050.

Supplemental Information

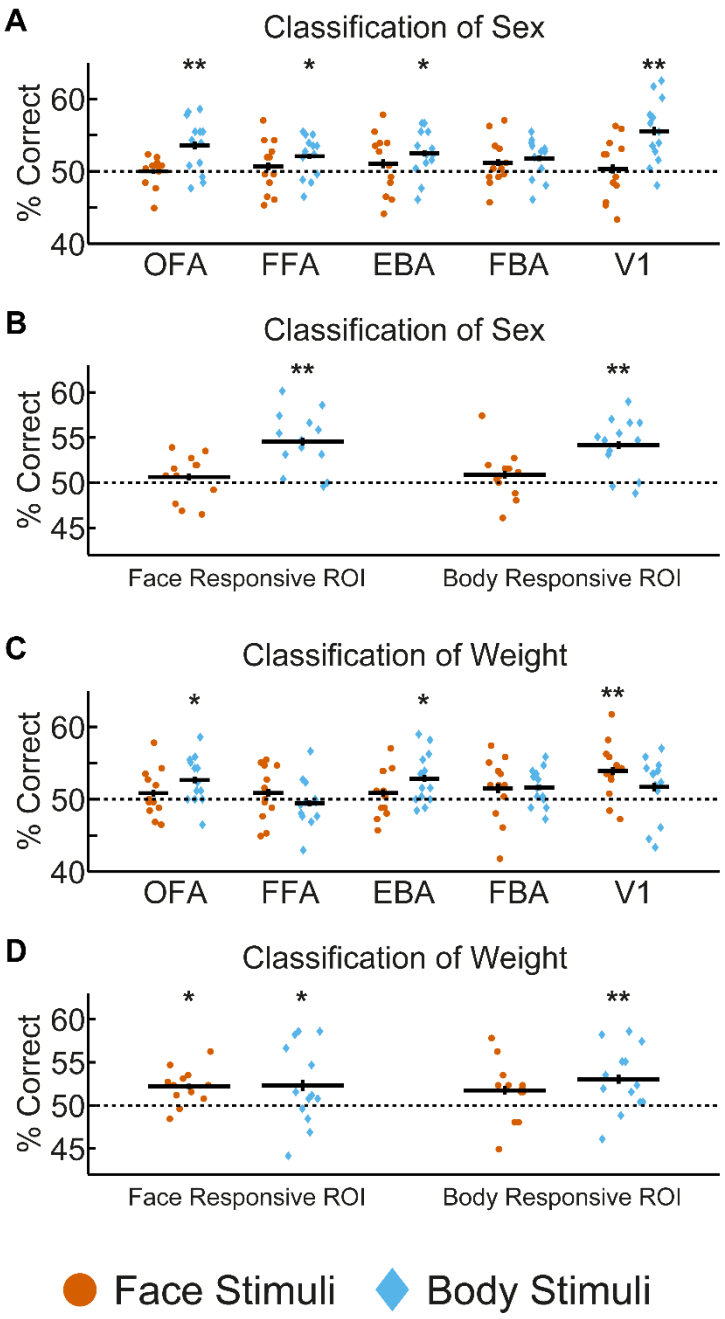

Figure S1. Classification of sex and weight in supplemental ROIs (face- and body-responsive ROIs defined at a constant t-contrast threshold, V1 defined as the 510 most posterior V1 voxels). Classification results for sex in separate ROIs (A) and distributed ROIs (B). Classification results for weight in separate ROIs (C) and distributed ROIs (D) Scatter plots show decoding accuracy for individual participants, horizontal black lines show group mean decoding accuracies and vertical error bars show the standard error of the mean. * indicates $p < 0.05$, ** indicates $p < 0.001$, Bonferroni-corrected for the number of ROIs (separate ROIs: N = 5; distributed ROIs: N = 2). Dotted lines indicate chance-level decoding performance, 50%.

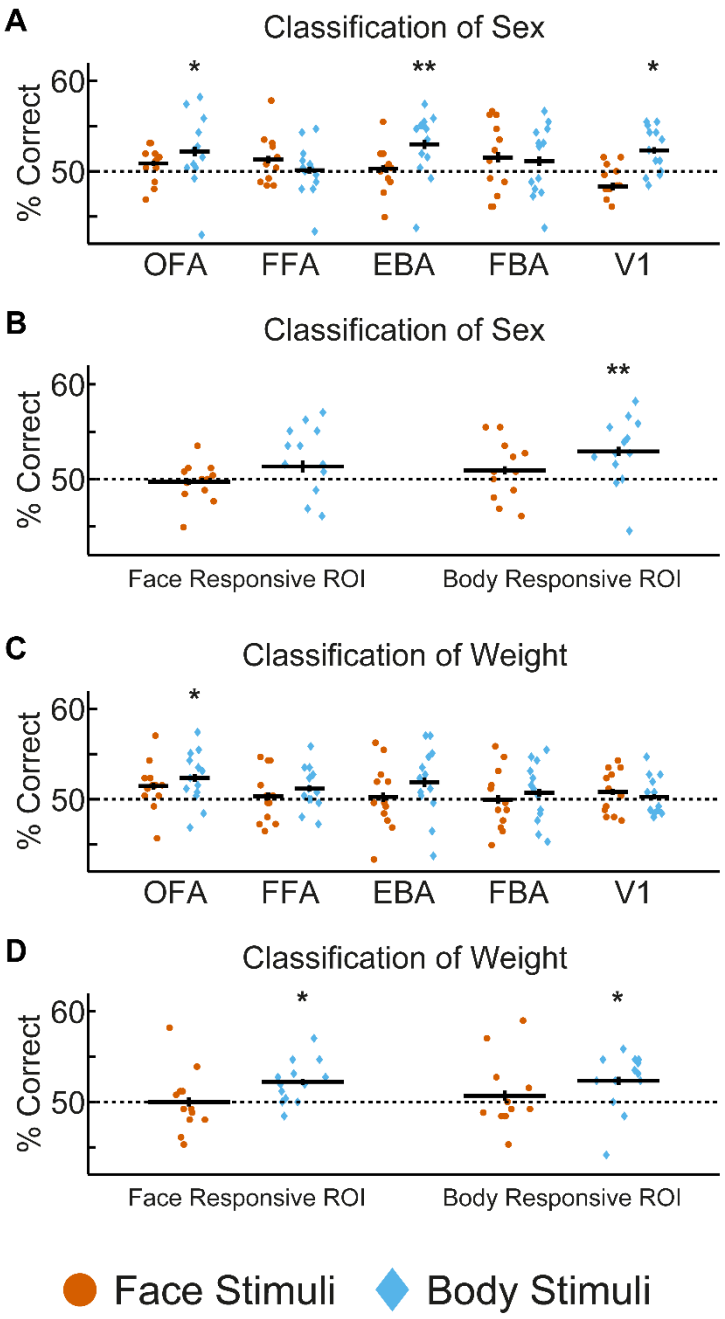

Figure S2. Size-invariant classification of sex and weight in supplemental ROIs (face- and body-responsive ROIs defined at a constant t-contrast threshold, V1 defined as the 510 most posterior V1 voxels). (A) and (B) show classification results for sex in separate ROIs (A) and distributed ROIs (B). (C) and (D) show classification results for weight in separate ROIs (C) and distributed ROIs (D). Scatter plots show decoding accuracy for individual participants, horizontal black lines show group mean decoding accuracies and vertical error bars show the standard error of the mean.* indicates $p < 0.05$, ** indicates $p < 0.001$, Bonferroni-corrected for the number of ROIs (separate ROIs: N = 5; distributed ROIs: N = 2). Dotted lines indicate chance-level decoding performance, 50%.

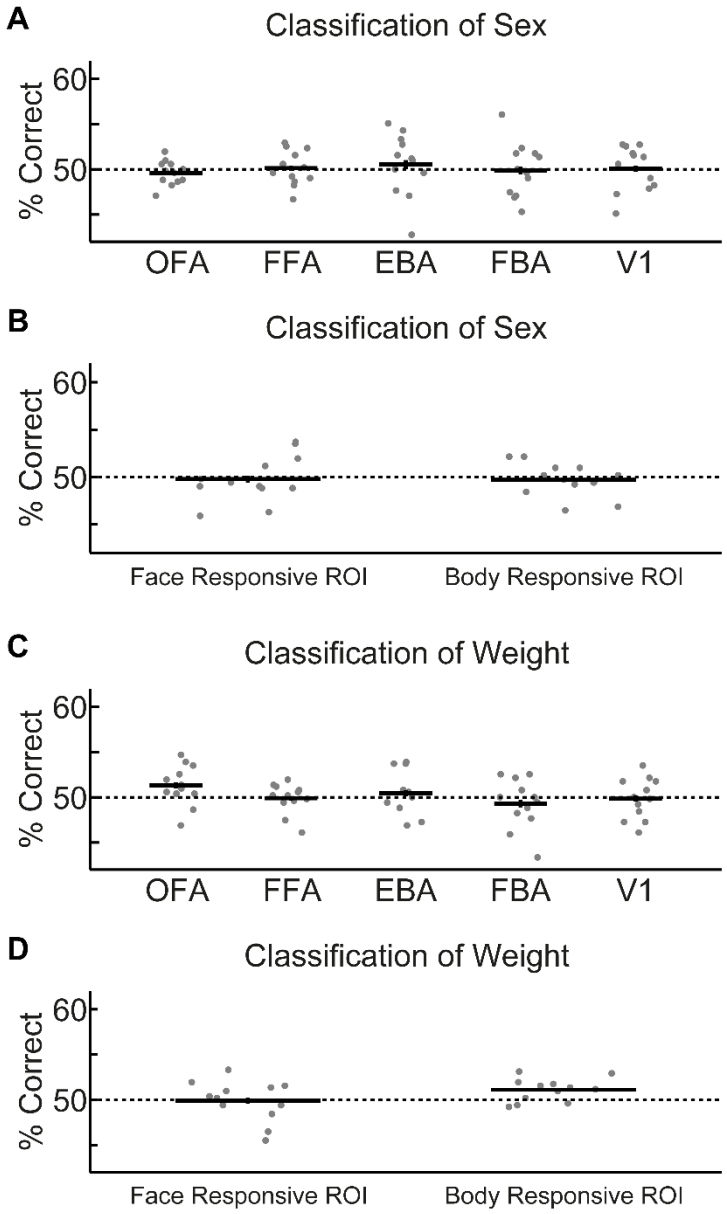

Figure S3. Classification of sex and weight across face and body stimuli in supplemental ROIs (face- and body-responsive ROIs defined at a constant t-contrast threshold, V1 defined as the 510 most posterior V1 voxels). In these cross-classification analyses, SVM classifiers were trained to distinguish sex and weight from neural activity when participants viewed faces and then subsequently tested on neural activity when participants viewed bodies, and vice-versa. (A) and (B) show classification results for sex in separate ROIs (A) and distributed ROIs (B). (C) and (D) show classification results for weight in separate ROIs (C) and distributed ROIs (D). Scatter plots show decoding accuracy for individual participants, horizontal black lines show group mean decoding accuracies and vertical error bars show the standard error of the mean. Dotted lines indicate chance-level decoding performance, 50%.

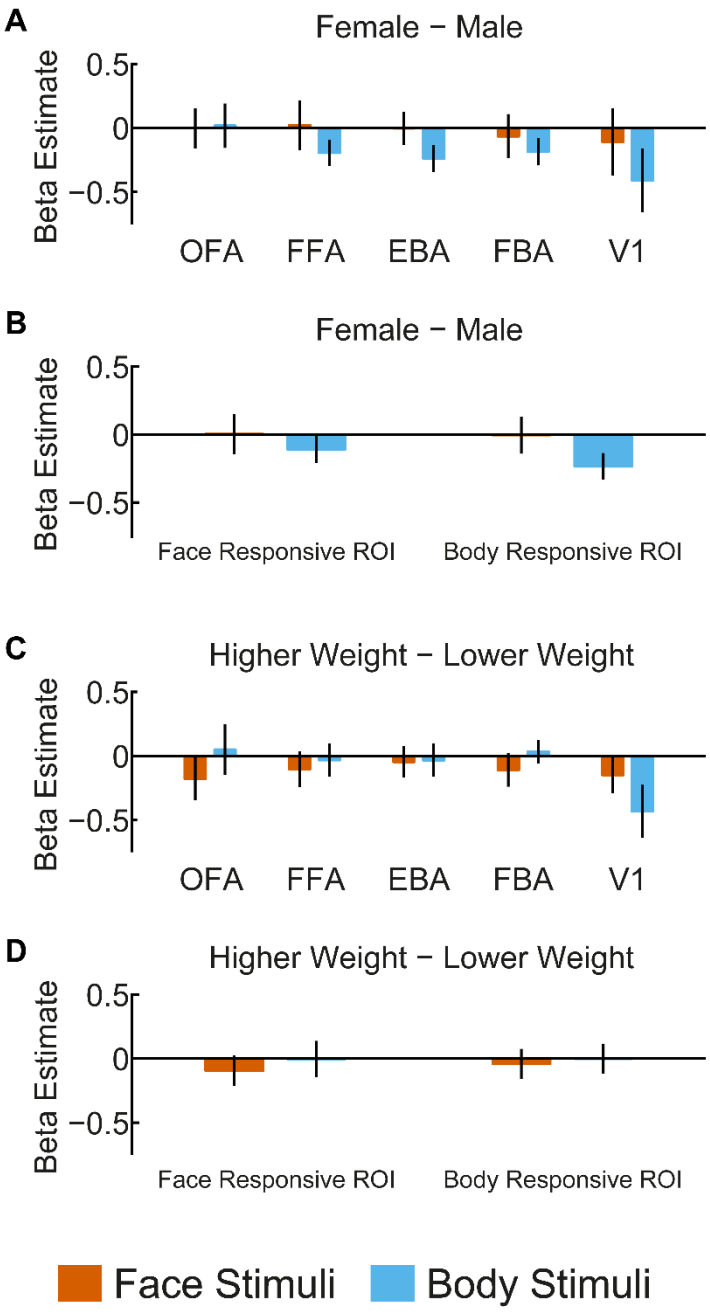

Figure S4. Mean BOLD differences between the categories sex and weight in supplemental ROIs (face- and body-responsive ROIs defined at a constant t-contrast threshold, V1 defined as the 510 most posterior V1 voxels). (A) and (B) illustrate mean differences between male and female stimuli for separate (A) and distributed (B) ROIs. Positive values indicate higher activation by female stimuli, negative values higher activation by male stimuli. (C) and (D) illustrate mean differences between higher and lower weight stimuli for separate (C) and distributed (D) ROIs. Positive values indicate higher activation by higher weight stimuli, negative values higher activation by lower weight stimuli. None of the differences were significant in any of the ROIs. Error bars show the standard error of the mean.

Figure S5. Mean BOLD differences between larger and smaller images in supplemental ROIs (face- and body-responsive ROIs defined at a constant t-contrast threshold, V1 defined as the 510 most posterior V1 voxels). (A) and (B) illustrate mean differences between larger and smaller image size stimuli for separate (A) and distributed (B) ROIs. Positive values indicate higher activation by larger size stimuli, negative values higher activation by smaller size stimuli. * indicates $p < 0.05$, ** indicates $p < 0.001$, Bonferroni-corrected for the number of ROIs (separate ROIs: N = 5; distributed ROIs: N = 2). Error bars show the standard error of the mean.

Table S1

ROI coordinates (in MNI space), ROI volume and number of subjects each ROI was identified in (N) for supplemental ROIs. Supplemental ROIs OFA, FFA, EBA and FBA were defined at a constant t-contrast threshold. Supplemental V1 was defined as the 510 most posterior V1 voxels in each participant. ROI analyses were conducted in subject space, and ROIs were normalized to MNI space in order to show the mean ROI locations. Coordinates show mean x, y and z locations and volume, ± standard deviations.

ROI	hemisphere	x	y	z	Volume (mm^3)	N
OFA	left	-34 ± 5.8	-83 ± 6.5	-13 ± 3.9	1630 ± 806.7	13
	right	38 ± 5.1	-79 ± 6.5	-13 ± 3.5	1873 ± 930.8	13
FFA	left	-41 ± 2.4	-49 ± 7.3	-21 ± 3.2	1595 ± 1106.2	13
	right	43 ± 3.2	-51 ± 5.8	-19 ± 2.2	2082 ± 495.6	13
EBA	left	-45 ± 2.7	-74 ± 4.6	5 ± 4.6	2390 ± 1123.9	13
	right	49 ± 2.5	-67 ± 3.1	0 ± 4.7	3447 ± 726.7	13
FBA	left	-42 ± 2.2	-50 ± 6.4	-21 ± 4.3	967 ± 679.2	12
	right	44 ± 2.5	-47 ± 5.2	-18 ± 2.9	1404 ± 732.1	13
V1	left	-9 ± 1.0	-89 ± 2.7	-1 ± 2.5	1864 ± 275.9	13
	right	10 ± 2.1	-86 ± 3.1	2 ± 2.0	2430 ± 356.3	13

Table S2

Comparison of results from ROI analyses conducted using ROIs defined at a set voxel size and ROIs defined at a constant t-contrast threshold. For each analysis, * indicates that classification was significant for both ROI definition methods.

Analysis		Separate ROIs					Distributed ROIs	
		OFA	FFA	EBA	FBA	V1	Face-responsive	Body-responsive
Classification	Sex of bodies	*	*	*		*	*	*
	Weight of bodies	*		*			*	*
	Sex of faces							
	Weight of faces					*		
Size-invariant classification	Sex of bodies			*				*
	Weight of bodies							*
	Sex of faces							
	Weight of faces							
Classification across face/body	Sex							
	Weight							

3. The neural coding of face and body identity

Abstract

Our visual system allows us to recognize the identity of a person across changes in viewpoint that substantially change the low level visual information reaching our retina. Previous neuroimaging studies have shown that the anterior face-responsive regions may be of particular importance for disentangling face identity from face viewpoint. Although it is known that we also use information from the body to identify people, much less is known about the brain regions involved in disentangling body identity information from body viewpoint, or where in the brain identity information from the face and body is combined. In this study, we trained participants to recognize three identities, and then recorded their brain activity using fMRI while they viewed images of the face and body (shown separately) of the three people from different viewpoints. Participants' task was to respond to the identity or viewpoint, allowing us to investigate if there would be differences in neural responses depending on whether participants attended to identity or viewpoint. We found consistent decoding of body identity across viewpoint in the fusiform body area (FBA), the right anterior temporal cortex and the middle frontal gyrus. This finding provides evidence of a similar importance of the fusiform and right anterior temporal cortex in disentangling identity from viewpoint for bodies as has previously been shown for faces, suggesting this is a general function of this area of cortex. In addition, we could decode identity in an abstract manner across neural activity evoked by faces and bodies in the early visual cortex, right inferior occipital cortex, right parahippocampal cortex and right superior parietal cortex, showing that several brain regions respond to person identity in an abstract manner. Lastly, we could decode identity more frequently when participants attended to identity, showing that participants' attention to identity enhances its neural representation.

Keywords: identity, face recognition, body recognition, invariance, viewpoint

3.1. Introduction

Our visual system allows us to recognize the identity of a person across changes in viewpoint, illumination, position, pose and expression. This is a remarkable ability as these changes lead to a great variability in low-level visual information arriving on the retina, yet we are able to distinguish between identities that look comparably similar to one another. It is not yet fully understood how the brain achieves this.

Face-responsive brain regions in the fusiform gyrus and anterior temporal cortex are thought to be important for our face recognition ability (Haxby, Hoffman, & Gobbini, 2000). These regions respond when participants recognize face identities (Grill-Spector, Knouf, & Kanwisher, 2004; Hoffman & Haxby, 2000; Nasr & Tootell, 2012) and dysfunction of these regions is associated with impairments in face recognition abilities (Barton, 2008; Busigny et al., 2014; Hadjikhani & de Gelder, 2002; Jonas et al., 2015). Anterior temporal cortex is thought to be of particular importance in encoding high-level face identity representations. Patterns of activity in this region can distinguish between different face identities (Kriegeskorte, Formisano, Sorger, & Goebel, 2007), and furthermore have been shown to be able to generalize across face viewpoint (Anzellotti, Fairhall, & Caramazza, 2014; Freiwald & Tsao, 2010; Guntupalli, Wheeler, & Gobbini, 2016), face expression (Nestor, Plaut, & Behrmann, 2011) and different halves of the same face (Anzellotti & Caramazza, 2015). Several studies have also shown that the fusiform face area (FFA) also responds to changes in identity (Andrews & Ewbank, 2004; Gauthier et al., 2000; Loffler, Yourganov, Wilkinson, & Wilson, 2005; Rotshtein, Henson, Treves, Driver, & Dolan, 2005; Winston, Henson, Fine-Goulden, & Dolan, 2004), and that face identity responses in the FFA can generalise across viewpoint (Anzellotti et al., 2014; Guntupalli et al., 2016). Some studies have also found high-level face identity responses in the occipital face area (OFA) (Anzellotti et al., 2014), the superior intraparietal sulcus (Jeong & Xu, 2016) and right inferior frontal cortex (Guntupalli et al., 2016).

Although psychological research has shown that we also use information from the body to recognize people (Hahn, O'Toole, & Phillips, 2015; O'Toole et al., 2011; Rice, Phillips, Natu, An, & O'Toole, 2013; Rice, Phillips, & O'Toole, 2013; Robbins & Coltheart, 2012), much less is known about the brain regions encoding body identity information. There is evidence

that the extrastriate and fusiform body areas (EBA and FBA) respond to body identity (Ewbank et al., 2011) and that the FBA as well as the inferior and medial frontal gyrus, cingulate gyrus, central and post-central sulcus and inferior parietal lobe respond more to the bodies of familiar people than to the bodies unfamiliar people (Hodzic, Kaas, Muckli, Stirn, & Singer, 2009). However, no study has investigated which regions of the human brain contain different patterns of neural activity evoked by different body identities, or furthermore which brain regions contain patterns of responses to different body identities that can generalize across different viewpoints. In macaques, electrophysiological recordings have shown that the macaque body-responsive patches contain body identity information that can generalize across viewpoint and pose (Kumar, Popivanov, & Vogels, 2017). Interestingly, identity decoding accuracy was higher in the more anterior body patch, suggesting there may be a similar importance of more anterior temporal regions in encoding body identity across viewpoint as has previously been found for faces in more anterior face-responsive regions. A first aim of the present study was to investigate which brain regions encode body identity in a viewpoint-invariant manner in the human brain.

A second aim of the present study was to investigate where identity information from the face and body is combined in the brain. It has been suggested that brain regions processing faces and bodies in occipitotemporal cortex are mostly separated, parallel networks (Pitcher, Charles, Devlin, Walsh, & Duchaine, 2009; Premereur, Taubert, Janssen, Vogels, & Vanduffel, 2016; Schwarzlose, Baker, & Kanwisher, 2005). However, a recent neuroimaging study found evidence of integration of face and body information in the EBA (Foster et al., 2019). In macaques, the anterior face patches have been shown to respond higher to images of a whole person compared to the addition of the responses to the face and body shown alone (Fisher & Freiwald, 2015), suggesting that these regions may integrate face and body information. A second aim of the present study was to investigate which brain regions contain similar patterns of response to a particular identity, regardless of whether it is viewed from an image of the face or the body.

In this study, we trained participants to recognize three identities and then recorded their brain activity using fMRI as they viewed images of the face and body of these three identities from three different viewpoints. Participants performed two behavioural tasks during the experiment, one where they responded to the stimulus identity and the other

where they responded to the stimulus viewpoint. This allowed us to investigate if there would be differences in participants' neural responses depending on which feature they attended to. First, we used linear support vector machine (SVM) classifiers to investigate which brain regions contain patterns of brain activity that could distinguish between the three face identities and between the three body identities. Second, we then further tested which brain regions contain patterns of brain activity evoked by the face or body identities that could generalize across viewpoint (e.g. a classifier was trained to distinguish between neural activity evoked by the face identities from two viewpoints and was then tested on distinguishing between the neural activity evoked by the face identities from the third viewpoint). Third, we tested which brain regions contain patterns of brain activity evoked by the identities that could generalize across neural activity evoked by faces and bodies (e.g. a classifier was trained to distinguish between neural activity evoked by the face identities and was then tested on distinguishing between neural activity evoked by the body identities). We performed all of these analyses in face- and body-responsive regions of interest (ROIs) as well as in whole-brain searchlight analyses, and performed them separately for fMRI data where participants performed the identity and viewpoint recognition behavioural tasks. These analyses allowed us to investigate which brain regions contain neural responses that can distinguish between the different face or body identities, which regions contain face or body identity responses that can generalize across viewpoint and which regions contain abstract identity responses that can generalize across neural activity evoked by the face and body. Furthermore, performing the analyses separately for the two behavioural tasks allowed us to investigate if identity decoding would be enhanced when participants' attended to the stimulus identity as compared to the stimulus viewpoint.

3.2. Materials and methods

This work was conducted using a dataset that was collected as part of a larger study. The data we present are novel analyses and results investigating neural and behavioural responses to face and body identity.

3.2.1. Participants

20 participants completed the experiment. One participant was excluded from the data analyses due to poor performance in the behavioural task (less than 40% correct responses in one condition). The remaining 19 participants (13 female, 21-51 years old) were included in the behavioural and fMRI analyses presented here. The experimental procedure was approved by the local ethics committee of the University Clinic Tübingen, and all participants provided written informed consent prior to the start of the experiment.

3.2.2. Stimuli

3.2.2.1. Main experiment stimuli

Our stimuli (Fig. 1A) consisted of separate face and body images of three identities from three viewpoints; 0° (front), 45° and 90° (profile). The three identities were all female, to ensure that sex did not differ between the three identities. For each identity, we recorded both a 3D face scan with a neutral expression and a 3D body scan in an A-pose. The face scans were then aligned to a 3D shape and expression model (Li, Bolkart, Black, Li, & Romero, 2017) and the body scans were aligned to a 3D shape and pose model (Loper, Mahmood, Romero, Pons-Moll, & Black, 2015). We then generated images of the three individuals from the three viewpoints (0°, 45° and 90°). The faces of the body images were covered using a grey rectangle in order to remove any face information from the body images.

For each identity, we also recorded a short video showing the whole body with the head fully visible turning between the left and right profile view. This video was used for identity learning prior to the fMRI experiment.

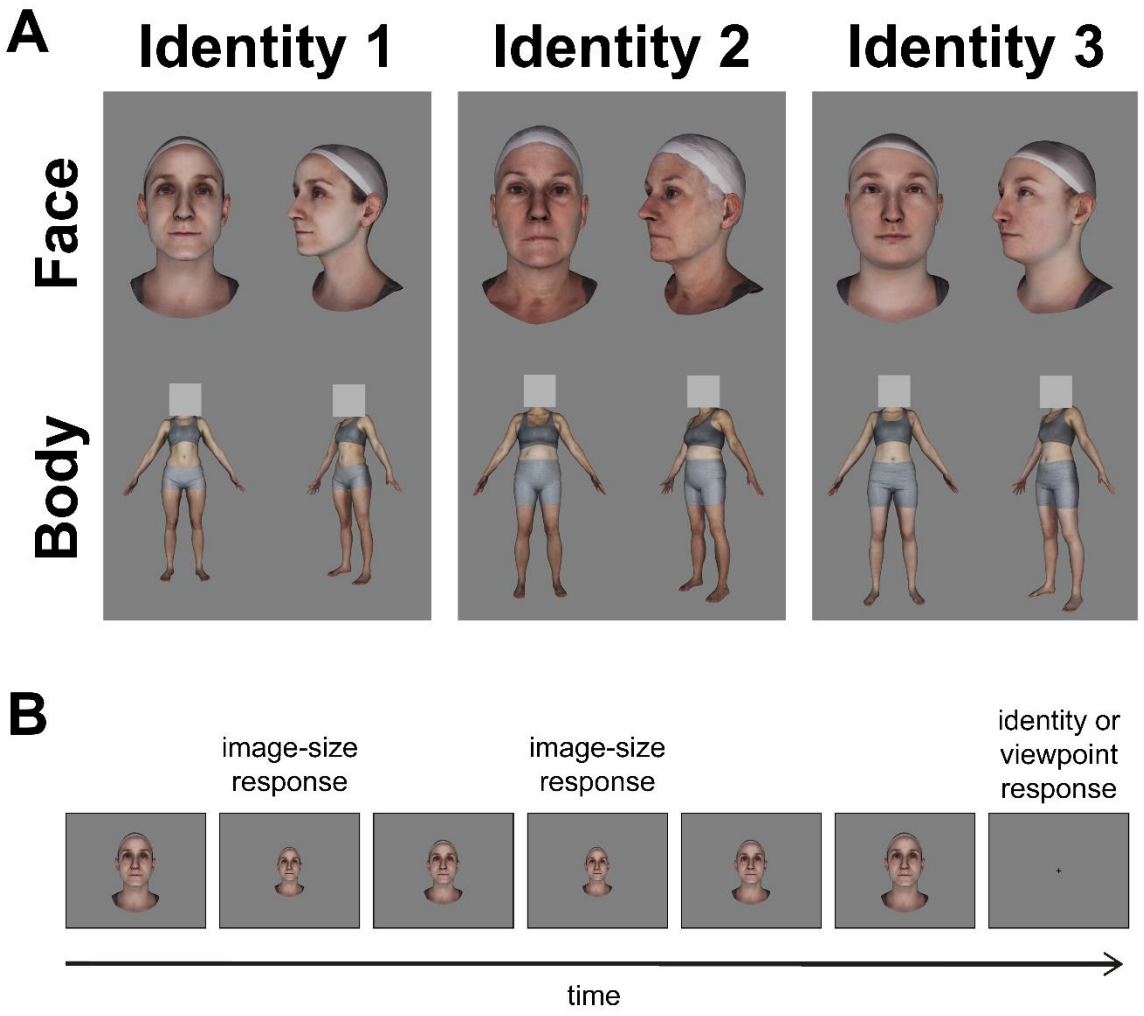

Figure 1. Experimental stimuli and procedure. (A) Stimuli were face and body images of three female identities shown from three viewpoints (0° and 45° shown here). (B) Example block of stimuli shown in the fMRI experiment. Participants viewed 6 images from one condition (i.e. face or body, one identity, one viewpoint) within a block, which varied in their image size (2 repetitions of 3 image sizes, shown in a random order). Participants performed two tasks; they responded immediately when they saw an image of the smallest image size, and they responded at the end of the block during fixation to indicate which identity or viewpoint was shown in the block (one half of experiment identity recognition task, other half viewpoint recognition task).

3.2.2.2. Localizer stimuli

Stimuli used to localize face- and body-responsive regions of interests were grayscale images of faces, headless bodies, objects and phase-scrambled images. The phase-scrambled images were generated by creating Fourier-scrambled versions of an image consisting of a collage of the face and headless body images.

3.2.3. Experimental procedure

The study consisted of a short identity learning session (outside of the MRI scanner) followed immediately by the main fMRI session, which consisted of eight runs of the main experiment and one run of a localizer experiment.

3.2.3.1. Identity learning procedure

Participants were trained to recognise the three stimulus identities from images of their face and body. The identity learning session consisted of five repetitions of a learning and testing with feedback procedure. During learning, participants viewed a 15 s video of each identity (showing their whole body turning between the left and right profile), then viewed the face and body images of the identity from the three viewpoints (0°, 45° and 90°), until the participant pressed a button to continue. A name was presented above each identity, so that participants could learn to associate each identity with its name. Following learning, participants completed 54 trials of the testing procedure with feedback. The 54 trials consisted of three repetitions of the stimulus conditions (face or body, three identities, three viewpoints) presented in a random order. In each trial, participants viewed a fixation cross for 1 s, then a stimulus image for 1 s, then a grey screen. Participants had up to 6 s to respond using a button press to indicate which identity was shown. After making a response, participants were given feedback as to whether their response was correct or not. At the end of the 54 trials, participants were shown an overall percentage correct score.

The identity learning session was presented on a laptop with resolution 1366x768, running Windows 10 with Matlab 2014a using the Psychophysics Toolbox extensions (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007; Pelli, 1997).

3.2.3.2. Main fMRI experiment procedure

Participants lay supine in the MRI scanner and viewed the stimuli on a screen positioned 92 cm behind their head, via a mirror attached to the head coil. Each run of the main experiment began with an instruction to the participant whether to respond to the identity or the viewpoint of the images in this run (4 runs each, see Section 3.2.3.2.1. for the task details). Participants viewed the experiment stimuli in a block design, where images within a block were from 1 of 18 conditions of a 2 (face or body) x 3 (identity) x 3 (viewpoint) factorial design. Each run contained 3 repetitions of all 18 conditions presented in a random order. The 18 conditions were preceded by and followed by 8 s of fixation.

Each block contained 6 images varying in their image size. There were 2 repetitions of 3 image sizes presented in a random order. The three image sizes had scale factors of 1, 1.3 and 1.6 (i.e. the largest image size was 1.6 times the width and height of the smallest image size). For face stimuli the mean widths and heights of the 3 image sizes were 4.4° x 6.4°, 3.6° x 5.2° and 2.8° x 4.0° of visual angle, for body stimuli the mean widths and heights of the 3 image sizes were 3.2° x 7.7°, 2.6° x 6.2° and 2.0° x 4.8° of visual angle. Each image was shown for 900 ms and a 100 ms blank screen was shown between images. Each block was followed by 2 s fixation.

The experiment was programmed with Matlab 2017a using the Psychophysics Toolbox extensions (Brainard, 1997; Kleiner et al., 2007) on Ubuntu 17.10. The experiment was presented using a projector with resolution 1920x1080 onto a screen with a width and height of 25° x 14° of visual angle.

3.2.3.2.1. Main fMRI experiment task

In half of the experiment runs participants were instructed to respond at the end of the block during fixation with a button press to indicate which identity was shown in the block (ID1, ID2 or ID3). In the other half of the experiment runs participants were instructed to respond at the end of the block during fixation to which viewpoint was shown in the block (0°, 45° or 90°).

In addition, participants were instructed to immediately press a button with their thumb whenever they saw an image that was the smallest of the three image sizes. This ensured participants kept their attention on the stimuli throughout each block.

3.2.3.3. fMRI localizer experiment procedure

Participants completed one run of a localizer experiment which was used to define face- and body-responsive brain regions. Participants viewed face, body, object and phase-scrambled images in a block design. Each block consisted of 8 images, which were each shown for 1.8s followed by a 0.2 s blank screen. Blocks were presented in a carryover counterbalanced sequence, such that face, body, object and phase scrambled blocks were preceded by each other block type an equal number of times (Brooks, 2012). Face, body and object images were shown in front of the phase-scrambled images to keep the area of retinal stimulation the same for all blocks. Participants performed a one-back matching task on the images to keep their attention on the stimuli. Images were repeated on average once every 9 s.

3.2.4. MRI sequence parameters

MRI data was acquired with a 3T Siemens Prisma scanner and a 64-channel head coil (Siemens, Erlangen, Germany). Functional T2* echoplanar images (EPI) were acquired using the following sequence parameters; multiband acceleration factor 2, GRAPPA acceleration factor 2, TR 1.84 s, TE 30 ms, flip angle 79°, FOV 192x192 mm. Volumes consisted of 60 slices and had an isotropic voxel size of 2x2x2 mm. We discarded the first 8 volumes of each run to allow for equilibration of the T1 signal. We additionally acquired a high-resolution T1-weighted anatomical scan for each participant with the following sequence parameters; TR 2 s, TE 3.06 ms, FOV 232x256 mm, 192 slices, isotropic voxel size of 1x1x1 mm.

3.2.5. MRI data preprocessing

We preprocessed our MRI data using SPM12 (http://www.fil.ion.ucl.ac.uk/spm/). Functional images were slice-time corrected, realigned and coregistered to the anatomical image. Functional images from the localizer experiment were additionally smoothed with a 6 mm Gaussian kernel. ROI and searchlight analyses on functional images from the main experiment were conducted on unsmoothed data in subject-space. The resulting searchlight

classification accuracy maps were then normalised to MNI (Montreal Neurological Institute) space, and spatially smoothed with a 6 mm Gaussian kernel. For the whole-brain univariate analyses the coregistered data was normalized to MNI space and spatially smoothed with a 6 mm Gaussian kernel.

3.2.6. Definition of regions of interest

Using fMRI data from the localizer experiment, we defined three face-responsive ROIs (the OFA, FFA and ATFA) and two body-responsive ROIs (the EBA and FBA). We first attempted to define the face-responsive ROIs using the contrast faces > objects and the body-responsive ROIs using the contrast bodies > objects. If we could not define a ROI in a participant using this contrast, we then attempted to define the ROI using the contrast faces > scrambled images or bodies > scrambled images. We initially used a contrast threshold of $p < 0.001$ (uncorrected) and reduced the threshold to $p < 0.01$ (uncorrected) if the ROI could not be defined with the initial threshold.

Table 1

Mean MNI coordinates and volume of each ROI, ± standard deviations. N shows the number of participants each ROI was identified in.

ROI	hem	x	y	z	Volume (mm^3)	N
OFA	left	-35 ± 6.9	-86 ± 5.9	-11 ± 3.6	731 ± 346.5	19
	right	38 ± 4.1	-81 ± 6.0	-10 ± 3.3	994 ± 382.8	19
FFA	left	-40 ± 2.8	-55 ± 5.5	-20 ± 2.8	709 ± 364.3	19
	right	42 ± 3.3	-52 ± 4.3	-18 ± 2.4	1083 ± 400.9	19
ATFA	left	-34 ± 5.5	-11 ± 6.7	-33 ± 6.9	177 ± 120.6	14
	right	34 ± 5.8	-8 ± 5.5	-37 ± 5.8	335 ± 265.6	18
EBA	left	-44 ± 3.7	-78 ± 5.4	3 ± 6.6	896 ± 486.0	19
	right	49 ± 2.3	-70 ± 2.7	0 ± 4.7	1686 ± 453.0	19
FBA	left	-39 ± 4.2	-50 ± 6.5	-20 ± 3.0	703 ± 459.0	18
	right	40 ± 3.9	-50 ± 5.6	-19 ± 2.4	1148 ± 552.6	19

3.2.7. Behavioural analyses

In half of the fMRI experiment runs, participants pressed a button to indicate the identity of the stimuli shown in each block, and in the other half of the runs they pressed a button to indicate the viewpoint of the stimuli shown in each block. We calculated participants' accuracy in identity and viewpoint detection using % correct. We performed behavioural analyses to investigate if there were any differences in participants' ability to detect the identity or viewpoint of the stimuli during the fMRI experiment, depending on the identity of the stimuli. To do this, we performed one-way repeated-measures ANOVAs with three levels (ID1, ID2 and ID3), separately for identity and viewpoint recognition of the face and body stimuli. Prior to each ANOVA, we performed a Mauchly's test of sphericity on the data and performed a Greenhouse-Geisser correction in any cases of non-sphericity.

3.2.8. Univariate fMRI analyses

We conducted univariate analyses to investigate if there were any differences in the mean BOLD signal evoked by the three stimulus identities. To do this, we used SPM12 to model the fMRI data with a GLM. The GLM contained regressors for each of the experimental conditions. We then performed one-way repeated-measures ANOVAs with three levels (ID1, ID2 and ID3) separately for face and body stimuli, in face- and body-responsive ROIs and in whole-brain analyses. For ROI analyses, we first performed a Mauchly's test of sphericity on the data and performed a Greenhouse-Geisser correction in any cases of non-sphericity. We then assessed significance using a threshold of $p < 0.05$, Bonferroni-corrected for N = 5 ROIs. Following any significant ANOVA results, we performed planned follow-up paired *t*-tests between the three identities to determine between which identities there were differences in neural activation. For whole-brain analyses, we assessed significance using a threshold of $p < 0.05$, false discovery rate (FDR) corrected.

3.2.9. Multivoxel pattern analyses (MVPA)

We conducted multivoxel pattern analyses (MVPA) to investigate if there were differences in the patterns of neural activity evoked by the three stimulus identities. To do this, we used SPM12 to model the fMRI data with a GLM. This GLM contained one regressor for each stimulus block. We then performed MVPA analyses on the beta weight images from the GLM using The Decoding Toolbox (Hebart, Görgen, & Haynes, 2015). We feature-scaled

the data using z-score normalisation, where we estimated the mean and standard deviation on the training data and applied these values to the test data. Any outlier values (greater than 2 standard deviations from the mean) were set to 2 or -2. We performed 3 different classification analyses (see Sections 3.2.9.1-3) using a linear SVM classifier (LIBSVM).

We performed all classification analyses in face- and body-responsive brain regions and in whole-brain searchlight analyses (4-voxel radius). For ROI analyses, significance was determined using permutation testing. Each analysis was repeated 10,000 times with the condition labels randomly assigned to generate a null distribution of mean classification accuracies expected by chance. We assessed significance by comparing how often we obtained a mean classification accuracy in the null distribution greater than or equal to the actual mean classification accuracy obtained for that ROI. We assessed significance using a threshold of $p < 0.05$, and used a Bonferroni-correction for N = 5 ROIs tested.

For searchlight analyses, we performed group analyses using nonparametric permutation tests with SnPM13 (http://warwick.ac.uk/snpm). We performed 10,000 permutations for each analysis and used 6 mm FWHM variance smoothing. We assessed significance with a threshold of $p < 0.05$, FDR corrected.

3.2.9.1. Identity classification analyses

We performed identity classification analyses to investigate which brain regions contain different patterns of activity evoked by different identities. We performed these analyses separately for neural activity evoked by face and body stimuli, and when participants performed the identity and viewpoint recognition tasks. We trained a linear SVM classifier to distinguish between patterns of neural activity evoked by the three identities using three runs of fMRI data. We then tested the classifier on its ability to predict the stimulus identities from neural activity in the fourth run of data. We performed a four-fold cross-validation procedure, where each run was used as the held out test dataset once, and we determined the final decoding accuracy by averaging over the four cross-validation iterations.

3.2.9.2. Viewpoint-invariant identity classification analyses

We performed viewpoint-invariant identity classification analyses to investigate which brain regions contain patterns of neural activity evoked by the stimulus identities that

can generalize across stimulus viewpoint. As previously, we performed these analyses separately for neural activity evoked by face and body stimuli, and when participants performed the identity and viewpoint recognition tasks. In these viewpoint-invariant analyses, we used three runs of fMRI data to train a linear SVM classifier to distinguish between patterns of neural activity evoked by the three identities from two of three viewpoints. We then tested the classifier on its ability to predict the stimulus identities from neural activity evoked by the third viewpoint in the fourth run of data. Again, we performed a four-fold cross-validation procedure (with each run used as the held out test dataset once) and also repeated the analysis three times with each viewpoint used as the held out test viewpoint once. We determined the final decoding accuracy by averaging over the four cross-validation iterations and the three viewpoint training and testing combinations.

3.2.9.3. Identity classification across face and body stimuli

We investigated which regions contain patterns of activity evoked by the stimulus identities that can generalize across neural activity evoked by faces and bodies. We performed these classification analyses separately for neural activity while participants performed the identity and viewpoint recognition tasks. We trained a linear SVM classifier to distinguish between patterns of neural activity evoked by the three face identities using three runs of fMRI data. We then tested the classifier on its ability to predict the identity of the body stimuli in the fourth run of data. We performed a four-fold cross-validation procedure (with each run used as the held out test dataset once) and also repeated the analysis using neural activity evoked by bodies for training the classifier and neural activity evoked by faces for testing it. We determined the final decoding accuracy by averaging over the four cross-validation iterations and the two training and test set combinations.

3.3. Results

3.3.1. Behavioural results

Participants responded using a button press at the end of each block to indicate the identity or viewpoint of the stimuli (half of the dataset for each task). We investigated if there were any differences in participants' identity or viewpoint recognition depending on the identity of the stimulus. The results are shown in Figure 2.

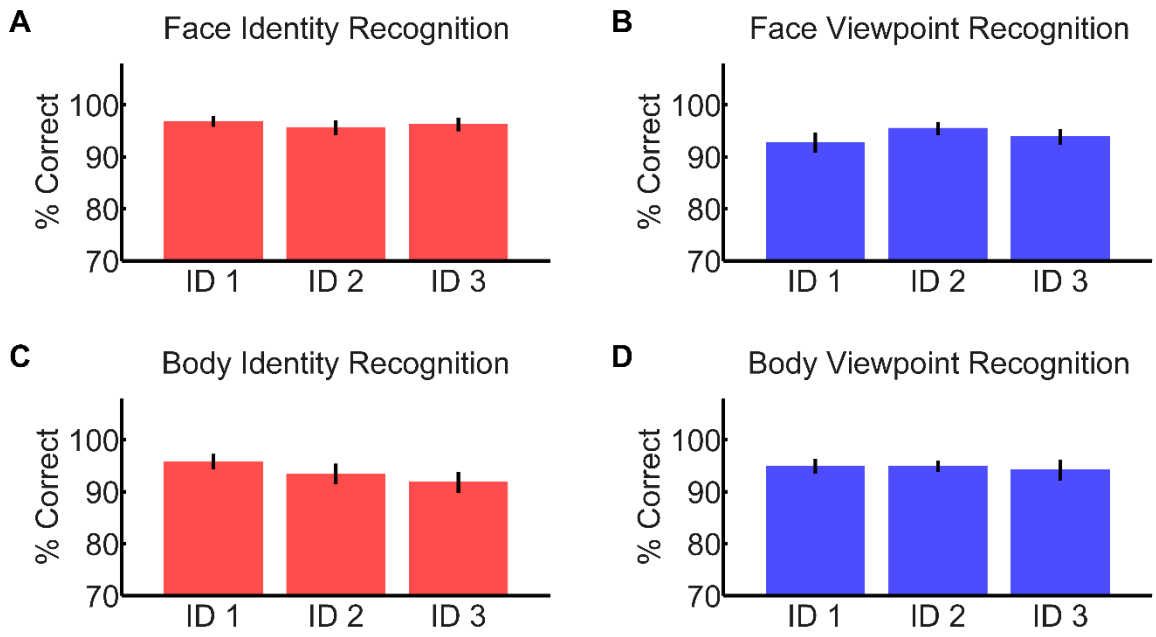

Figure 2. Recognition of the identity and viewpoint of the three stimulus identities (ID1, ID2 & ID3). (A) and (C) show identity recognition accuracy (% correct) for the three stimulus identities from the face (A) and body (C) images. (B) and (D) show viewpoint recognition accuracy (% correct) for the three stimulus identities from the face (B) and body (D) images. Error bars indicate ±1 SEM.

3.3.1.1. Identity recognition

Participants showed high identity recognition performance for both face (96.2 %) and body (93.7 %) stimuli. We investigated if there were any differences in our participants' ability to recognise the three identities. One-way repeated measures ANOVAs with three levels (ID1, ID2 and ID3) showed that there were no significant differences in participants'

ability to recognise the three identities from the face ($F_{2,36}$ = 0.53, p = 0.59, η_p^2 = 0.029) or body ($F_{2,36}$ = 1.20, p = 0.31, η_p^2 = 0.063) stimuli. These results show that participants could easily recognise all stimuli identities from the face and body.

3.3.1.2. Viewpoint recognition

Participants showed high viewpoint recognition performance for both face (94.0 %) and body (94.6 %) stimuli. We investigated if there were any differences in participants' viewpoint recognition performance depending on the identity of the stimulus. One-way repeated measures ANOVAs with three levels (ID1, ID2 and ID3) showed that there were no significant differences in participants' ability to recognise the viewpoint of the stimuli between the three face identities ($F_{2,36}$ = 2.04, p = 0.14, η_p^2 = 0.10) or between the three body identities ($F_{2,36}$ = 0.18, p = 0.84, η_p^2 = 0.010). These results show that participants could recognise the stimulus viewpoint equally well regardless of the stimulus identity.

3.3.2. Univariate fMRI results

We investigated whether there were any differences in the mean BOLD activity evoked by the three identities. To do this, we performed one-way repeated measures ANOVAs with 3 levels (ID1, ID2 and ID3) in face- and body-responsive ROIs and in whole-brain analyses. The results are shown in Figure 3.

3.3.2.1. Face identity responses

We performed one-way repeated measures ANOVAs with 3 levels (ID1, ID2 and ID3) to test whether there were any differences in the mean BOLD activity evoked by the three face identities. For fMRI data from the identity recognition task, we found no significant differences between the mean BOLD activity evoked by the three face identities in any of our face- or body-responsive ROIs (Fig. 3A; OFA: $F_{2,36}$ = 6.53, p = 0.011 uncorrected for multiple comparisons and Greenhouse-Geisser corrected for non-sphericity, η_p^2 = 0.27; FFA: $F_{2,36}$ = 3.52, p = 0.040 uncorrected, η_p^2 = 0.16; ATFA: $F_{2,36}$ = 1.52, p = 0.23 uncorrected, η_p^2 = 0.078; EBA: $F_{2,36}$ = 4.51, p = 0.018 uncorrected, η_p^2 = 0.20; FBA: $F_{2,36}$ = 1.71, p = 0.20 uncorrected, η_p^2 = 0.087) or in any other region in a whole-brain analysis.

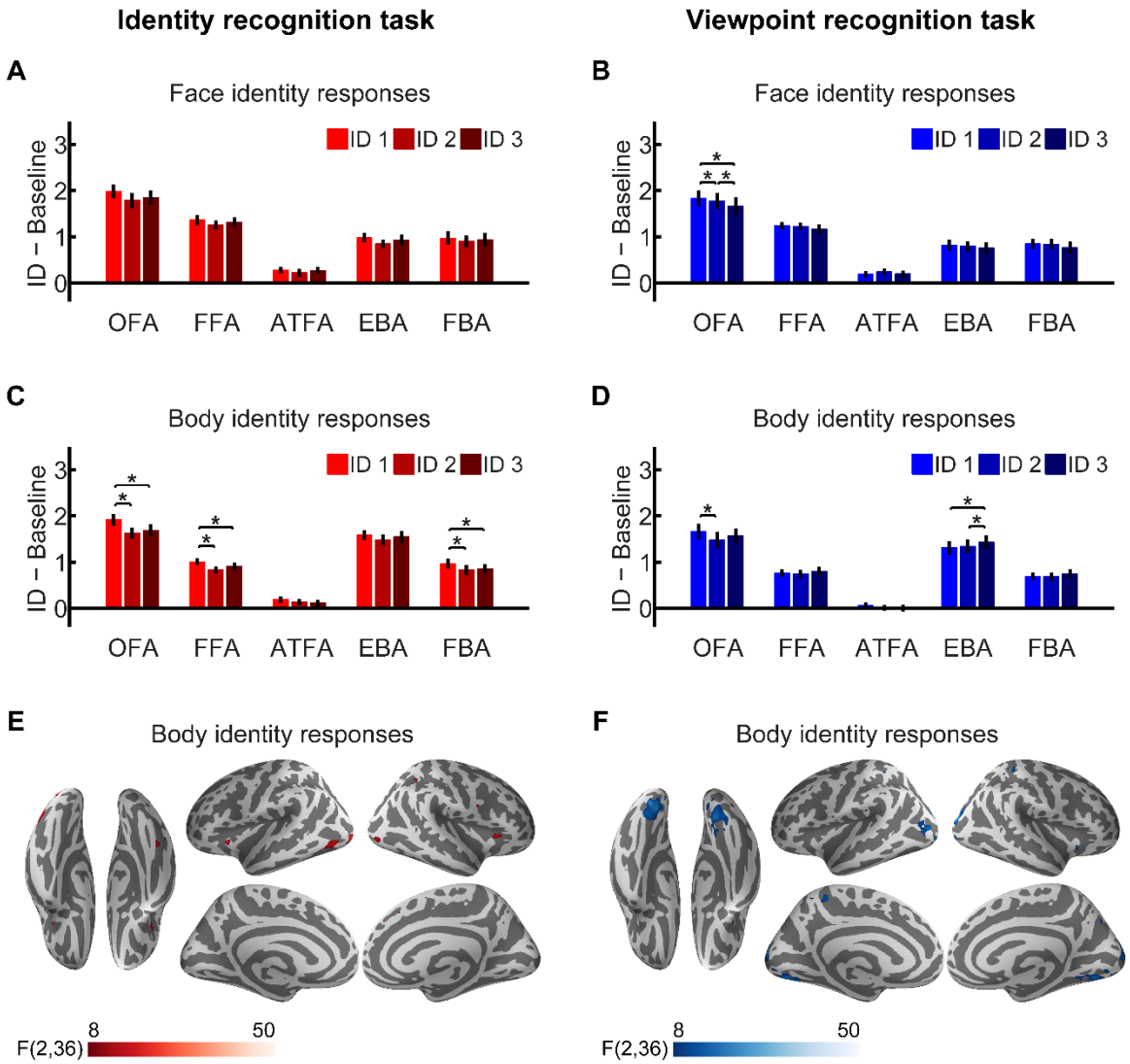

Figure 3. Differences in mean BOLD response to the three identities. (A) and (B) show mean BOLD responses to the three face identities in face- and body-responsive ROIs during the identity (A) and viewpoint (B) recognition task. * indicates $p < 0.05$. (C) and (D) show mean BOLD responses to the three body identities in face- and body-responsive ROIs during the identity (C) and viewpoint (D) recognition task. * indicates $p < 0.05$ (E) and (F) show differences in mean BOLD responses to the three body identities in whole-brain analyses (FDR corrected) during the identity (E) and viewpoint (F) recognition task.

For fMRI data from the viewpoint recognition task (Fig. 3B), we found significant differences between the mean BOLD activity evoked by the three face identities in the OFA ($F_{2,36}$ = 10.27, p = 0.0075 Bonferroni corrected and Greenhouse-Geisser corrected for non-sphericity, η_p^2 = 0.36) but not in any other ROIs (FFA: $F_{2,36}$ = 2.47, p = 0.12 uncorrected for multiple comparisons and Greenhouse-Geisser corrected for non-sphericity, η_p^2 = 0.12; ATFA: $F_{2,36}$ = 0.99, p = 0.38 uncorrected, η_p^2 = 0.052; EBA: $F_{2,36}$ = 1.55, p = 0.23 uncorrected, η_p^2 = 0.079; FBA: $F_{2,36}$ = 2.70, p = 0.097 uncorrected for multiple comparisons and Greenhouse-Geisser corrected for non-sphericity, η_p^2 = 0.13). Follow-up paired *t*-tests showed that in the OFA there was higher activity to ID1 compared to ID2 (M = 0.056, SE = 0.023, t_{18} = 2.43, p = 0.026, Cohen's d = 0.56) and ID3 (M = 0.17, SE = 0.042, t_{18} = 4.03, p = 7.78 x 10^{-4}, Cohen's d = 0.93), and higher activity to ID2 than ID3 (M = 0.11, SE = 0.045, t_{18} = 2.50, p = 0.023, Cohen's d = 0.57). In addition, a whole-brain analysis identified small, bilateral clusters in the early visual cortex showing differences in BOLD activity to the three face identities during the viewpoint recognition task.

3.3.2.2. Body identity responses

We tested whether there were any differences in the mean BOLD activity evoked by the three body identities using one-way repeated measures ANOVAs with 3 levels (ID1, ID2 and ID3). For fMRI data from the identity recognition task (Fig. 3C), we found significant differences in the mean BOLD activity evoked by the three body identities in the FBA ($F_{2,36}$ = 6.96, p = 0.014 Bonferroni corrected, η_p^2 = 0.28), OFA ($F_{2,36}$ = 20.76, p = 5.04 x 10^{-6} Bonferroni corrected, η_p^2 = 0.54) and FFA ($F_{2,36}$ = 11.21, p = 8.21 x 10^{-4} Bonferroni corrected, η_p^2 = 0.38), but not in the EBA ($F_{2,36}$ = 2.02, p = 0.15 uncorrected, η_p^2 = 0.10) or ATFA ($F_{2,36}$ = 1.75, p = 0.19 uncorrected, η_p^2 = 0.089). Follow-up paired *t*-tests in the FBA, OFA and FFA showed that in all three ROIs there was higher activity to ID1 compared to ID2 (FBA: M = 0.14, SE = 0.028, t_{18} = 5.03, p = 8.75 x 10^{-5}, Cohen's d = 1.15; OFA: M = 0.29, SE = 0.044, t_{18} = 6.56, p = 3.66 x 10^{-6}, Cohen's d = 1.50; FFA: M = 0.17, SE = 0.030, t_{18} = 5.67, p = 2.25 x 10^{-5}, Cohen's d = 1.30) and ID3 (FBA: M = 0.11, SE = 0.047, t_{18} = 2.40, p = 0.027, Cohen's d = 0.55; OFA: M = 0.23, SE = 0.046, t_{18} = 4.98, p = 9.73 x 10^{-5}, Cohen's d = 1.14; FFA: M = 0.10, SE = 0.038, t_{18} = 2.67, p = 0.016, Cohen's d = 0.61) but no difference between activity to ID2 and ID3 (FBA: M = -0.028, SE = 0.043, t_{18} = -0.66, p = 0.52, Cohen's d = -0.15; OFA: M = -0.058, SE = 0.051, t_{18} = -1.13, p = 0.27, Cohen's d = -0.26; FFA: M = -0.071, SE = 0.041, t_{18} = -1.74, p =

0.098, Cohen's *d* = -0.40). We performed a whole-brain analysis to investigate if there were any additional regions showing differences in mean BOLD activity to the three body identities during the identity recognition task (Fig. 3E). We identified bilateral clusters in the early visual cortex, occipitotemporal cortex (overlapping with the locations of the OFA, FFA and FBA) and insula cortex, and unilateral clusters in the right inferior parietal cortex, right precuneus and right medial superior frontal gyrus.

For fMRI data from the viewpoint recognition task (Fig. 3D), we found significant differences in activity evoked by the three body identities in the OFA ($F_{2,36}$ = 6.52, p = 0.019 Bonferroni corrected, η_p^2 = 0.27) and EBA ($F_{2,36}$ = 6.11, p = 0.026 Bonferroni corrected, η_p^2 = 0.25), but not in any other ROIs (FBA: $F_{2,36}$ = 1.88, p = 0.17 uncorrected, η_p^2 = 0.095; FFA: $F_{2,36}$ = 1.43, p = 0.25 uncorrected, η_p^2 = 0.074; ATFA: $F_{2,36}$ = 2.27, p = 0.12 uncorrected, η_p^2 = 0.11). Follow-up paired *t*-tests in the OFA showed there was lower activity to ID2 compared to ID1 (M = -0.19, SE = 0.053, t_{18} = -3.55, p = 0.0023, Cohen's d = -0.81). Follow-up paired *t*-tests in the EBA showed there was higher activity to ID3 compared to ID1 (M = 0.12, SE = 0.036, t_{18} = 3.43, p = 0.0030, Cohen's d = 0.79) and ID2 (M = 0.091, SE = 0.036, t_{18} = 2.51, p = 0.022, Cohen's d = 0.58). We performed a whole-brain analysis to investigate if any other regions would show different levels of mean response to the three body identities during the viewpoint recognition task (Fig. 3F). We identified bilateral clusters in the early visual cortex, middle occipital cortex, fusiform gyrus, superior temporal cortex, superior parietal cortex, precuneus, superior frontal cortex and insula cortex.

3.3.3. Face identity MVPA

We investigated which brain regions contain separable patterns of neural responses evoked by the three face identities. We performed multivoxel pattern analyses to investigate which regions could classify face identity, and classify face identity across viewpoint, from patterns of neural activity. We performed all analyses in face- and body-responsive ROIs and in whole-brain searchlight analyses. The results are shown in Figure 4.

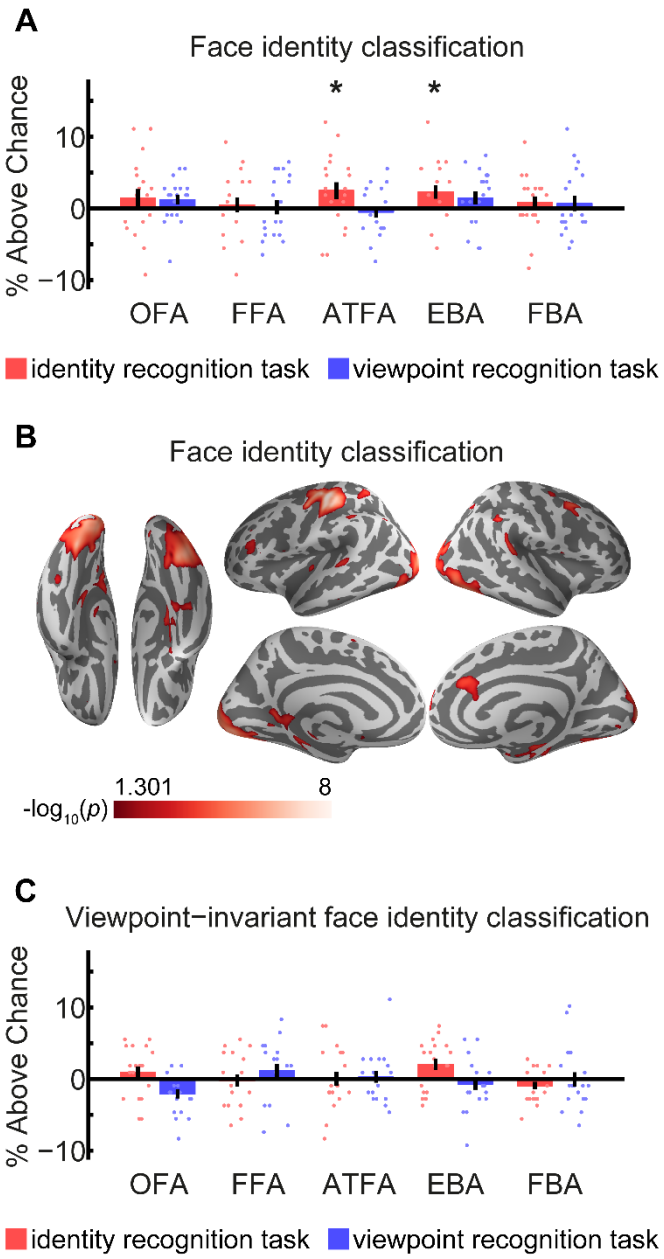

Figure 4. Classification and viewpoint-invariant classification of face identity. (A) shows face identity classification in face- and body-responsive ROIs. Scatter points show classification accuracies for individual participants, error bars show ±1 SEM, * indicates $p < 0.05$ Bonferroni corrected. (B) shows classification of face identity during the identity recognition task in a whole-brain searchlight analysis. The scale bar shows $-\log_{10}(p$ values) between 1.301 ($p = 0.05$) and 8 ($p = 1 \times 10^{-8}$), FDR corrected. (C) shows viewpoint-invariant face identity classification in face- and body-responsive ROIs. Scatter points show classification accuracies for individual participants, error bars show ±1 SEM.

3.3.3.1. Face identity classification

We investigated which regions could classify face identity by training a linear SVM classifier to distinguish between patterns of neural activity evoked by the three face identities, using three runs of fMRI data. We then tested the classifier on its ability to decode face identity from the fourth run. We used a four-fold cross-validation procedure, such that each run was used as the held-out test run once. We performed the analysis twice separately; once using fMRI data where participants performed the identity recognition task, once using fMRI data where participants performed the viewpoint recognition task.

We first performed a ROI analysis (Fig. 4A) to investigate which face- and body responsive ROIs could decode face identity above chance-level (33 ⅓ %). Using fMRI data from the identity recognition task, we were able to decode face identity significantly above chance from the face-responsive ATFA (35.8 %, p = 0.031 Bonferroni corrected, Cohen's d = 0.49) and body-responsive EBA (35.6 %, p = 0.045 Bonferroni corrected, Cohen's d = 0.56), but not from any other face-responsive ROI (OFA: 34.8 %, p = 0.064 uncorrected, Cohen's d = 0.27; FFA: 33.8 %, p = 0.31 uncorrected, Cohen's d = 0.11) or the FBA (34.1 %, p = 0.20 uncorrected, Cohen's d = 0.20). Using fMRI data from the viewpoint recognition task, we were not able to decode face identity from the ATFA (32.8 %, p = 0.70 uncorrected, Cohen's d = -0.16), EBA (34.8 %, p = 0.067 uncorrected, Cohen's d = 0.37) or any other ROI (OFA: 34.6 %, p = 0.10 uncorrected, Cohen's d = 0.40; FFA: 33.5 %, p = 0.44 uncorrected, Cohen's d = 0.034; FBA: 34.0 %, p = 0.22 uncorrected, Cohen's d = 0.15).

Secondly, we performed a whole-brain searchlight analysis to investigate if any other brain regions could decode face identity. Using fMRI data from the identity recognition task (Fig. 4B) we identified clusters than could decode face identity bilaterally in the early visual cortex, inferior occipital cortex, fusiform gyrus, superior parietal cortex, superior temporal cortex and parahippocampal gyrus, and unilaterally in the right middle frontal gyrus, right anterior cingulum, right medial superior frontal gyrus and left inferior frontal gyrus. We could also decode identity in the left motor cortex as participants pressed different buttons for each stimulus identity. Using fMRI data from the viewpoint recognition task, we were unable to decode face identity from any regions.

3.3.3.2. Viewpoint-invariant face identity classification

We next investigated which regions could classify face identity across viewpoint. To do this, we trained a linear SVM classifier to distinguish between patterns of neural activity evoked by the three face identities using two of the three stimulus viewpoints and three runs of fMRI data. We then tested the classifier on its ability to decode face identity from the third stimulus viewpoint in the fourth run of fMRI data. Again, we used a four-fold cross-validation with each run used as the test set once and, in addition, we repeated the analysis three times using each viewpoint as the left out test viewpoint once. As previously, we performed the analysis twice separately using data from the identity and viewpoint recognition tasks.

We were unable to decode face identity across viewpoint from any of the ROIs we tested (Fig. 4C) using fMRI data from the identity recognition task (OFA: 34.3 %, p = 0.17 uncorrected, Cohen's d = 0.26; FFA: 33.1 %, p = 0.62 uncorrected, Cohen's d = -0.066; ATFA: 33.4 % p = 0.49 uncorrected, Cohen's d = 0.012; EBA: 35.4 %, p = 0.012 uncorrected, Cohen's d = 0.60; FBA: 32.4 %, p = 0.86 uncorrected, Cohen's d = -0.40) or the viewpoint recognition task (OFA: 31.2 %, p = 0.99 uncorrected, Cohen's d = -0.77; FFA: 34.5 %, p = 0.11 uncorrected, Cohen's d = 0.27; ATFA: 33.6 % p = 0.37 uncorrected, Cohen's d = 0.083; EBA: 32.7 %, p = 0.77 uncorrected, Cohen's d = -0.19; FBA: 33.3 %, p = 0.51 uncorrected, Cohen's d = -0.011).

We performed searchlight analyses to investigate if any other brain regions would be able to decode face identity across viewpoint. We did not identify any regions in these analyses.

3.3.4. Body identity MVPA

We investigated which brain regions contain separable patterns of neural responses evoked by the three body identities. To do this, we performed multivoxel pattern analyses to investigate which regions could classify body identity, and classify body identity across viewpoint. We performed these analyses in face- and body-responsive ROIs (Fig. 5) and in whole-brain searchlight analyses (Fig. 6).

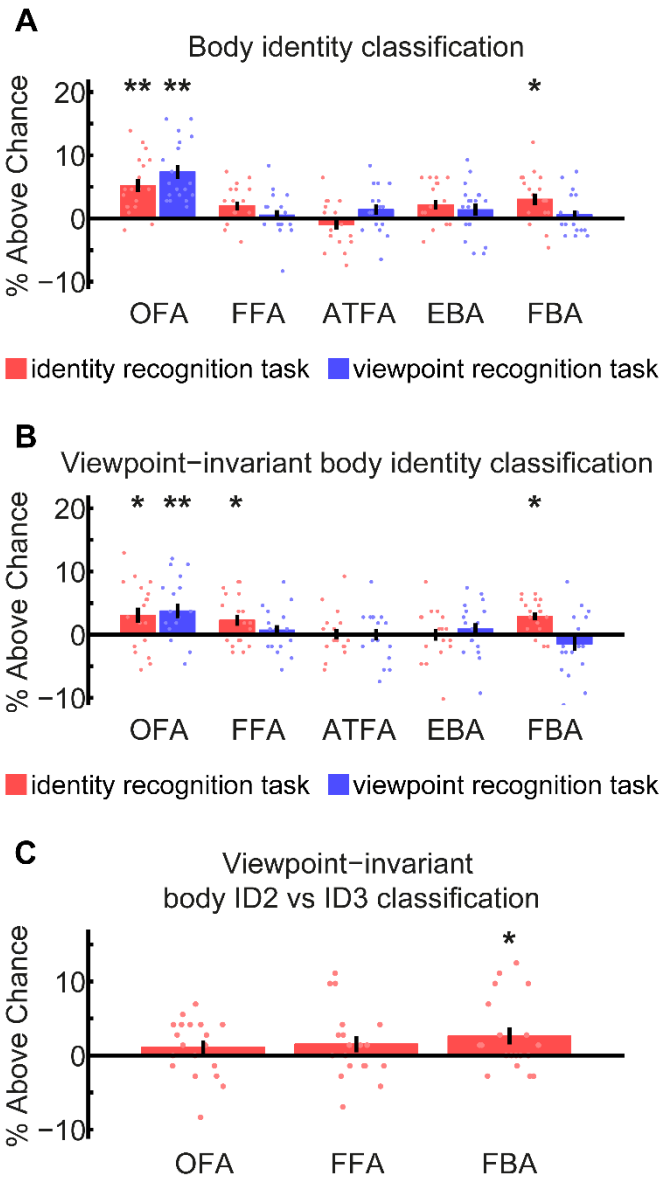

Figure 5. Classification and viewpoint-invariant classification of body identity in face- and body-responsive ROIs. (A) shows body identity classification, (B) shows viewpoint-invariant body identity classification and (C) shows viewpoint-invariant body identity classification for ID2 vs. ID3 only, with fMRI data from the identity recognition task. Scatter points show classification accuracies for individual participants and error bars show ±1 SEM. ** indicates $p < 0.001$, * indicates $p < 0.05$, Bonferroni corrected.

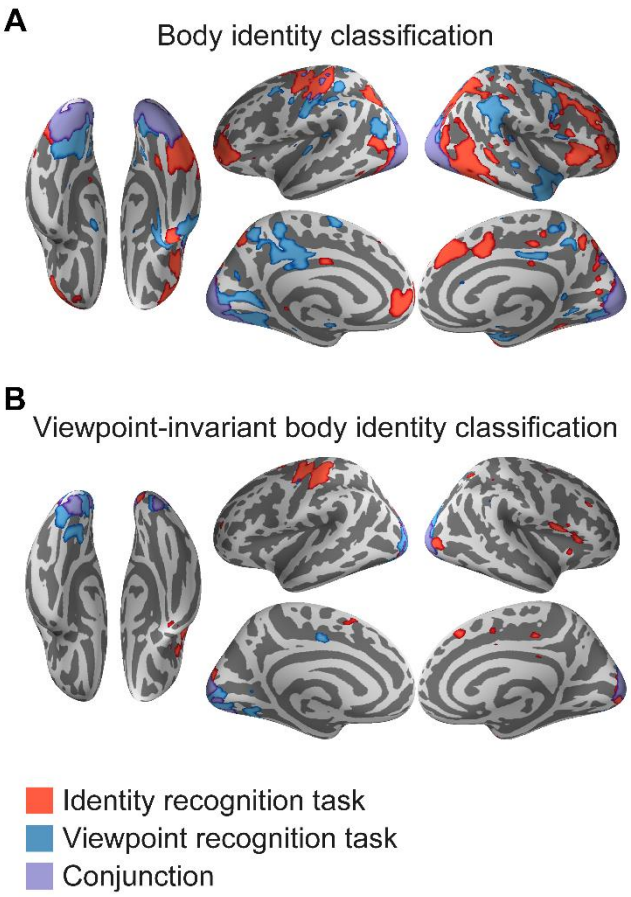

Figure 6. Classification (A) and viewpoint-invariant classification (B) of body identity in whole-brain searchlight analyses. Regions showing significant activity during the identity recognition task are shown in red, regions showing significant activity during the viewpoint recognition task are shown in blue and regions showing significant activity during both tasks (conjunction) are shown in purple. Significant regions were defined using a $p < 0.05$ FDR correction.

3.3.4.1. Body identity classification

To investigate which brain regions contain separable patterns of neural activity evoked by different body identities, we first trained a linear SVM to distinguish between patterns of neural activity evoked by the three body identities, using three runs of fMRI data. We then tested the trained classifier on its ability to decode body identity from the fourth run. Again, we used a four-fold cross-validation procedure and performed the analysis twice separately using fMRI data recorded while participants performed the identity and viewpoint recognition tasks.

First, we investigated which of our face- and body-responsive ROIs (Fig. 5A) could decode body identity above chance-level (33 ⅓ %). Using fMRI data from the identity recognition task, we could decode body identity significantly above chance from the body-responsive FBA (36.4 %, p = 0.0045 Bonferroni corrected, Cohen's d = 0.76) and face-responsive OFA (38.5 %, p < 0.0006 Bonferroni corrected, Cohen's d = 1.18), but not from the body-responsive EBA (35.5 %, p = 0.014 uncorrected, Cohen's d = 0.65) or any other ROIs (FFA: 35.3 %, p = 0.018 uncorrected, Cohen's d = 0.67; ATFA: 32.4 %, p = 0.84 uncorrected, Cohen's d = -0.28). Using fMRI data from the viewpoint recognition task, we were able to decode body identity from the OFA (40.7 %, p < 0.0006 Bonferroni corrected, Cohen's d = 1.54), but not from any other ROIs (EBA: 34.7 %, p = 0.078 uncorrected, Cohen's d = 0.33; FBA: 33.9 %, p = 0.27 uncorrected, Cohen's d = 0.19; FFA: 33.9 %, p = 0.28 uncorrected, Cohen's d = 0.16; ATFA: 34.7 %, p = 0.069 uncorrected, Cohen's d = 0.39).

Second, we performed a whole-brain searchlight analysis to investigate if we could decode body identity from any other brain regions (Fig. 6A). We could decode body identity from a large area of occipital cortex using fMRI data from both the identity and viewpoint recognition tasks. Using fMRI data from the identity recognition task, we could also decode body identity from bilateral regions in the fusiform gyrus, superior parietal cortex, inferior frontal gyrus and middle frontal gyrus, and unilaterally from the right anterior temporal cortex and right insula cortex. We could also decode body identity in the left motor cortex as participants pressed different buttons to indicate the stimulus identity. Using fMRI data from the viewpoint recognition task, we could also decode body identity from bilateral regions in the fusiform gyrus, superior parietal cortex, supramarginal gyrus, cingulum, precentral gyrus and the caudate nucleus, and unilaterally from the right superior frontal gyrus.

3.3.4.2. Viewpoint-invariant body identity classification

Next, we investigated which brain regions encode body identity in a viewpoint-invariant manner. To do this, we trained a linear SVM classifier to distinguish between patterns of neural activity evoked by the three body identities from two viewpoints, using three runs of fMRI data. We then tested the trained classifier on its ability to decode body identity from the third viewpoint in the fourth fMRI run. We used a four-fold cross-validation with each run used as the test set once and we repeated the analysis three times

using each viewpoint as the left out test viewpoint once. As previously, we performed the analysis twice separately using data from the identity and viewpoint recognition tasks.

We first tested which of our face- and body-responsive ROIs could decode body identity across viewpoint (Fig. 5B). Using fMRI data from the identity recognition task, we could decode body identity across viewpoint significantly above chance-level (33 ⅓ %) in the body-responsive FBA (36.2 %, p = 0.0030 Bonferroni corrected, Cohen's d = 1.06) and face-responsive OFA (36.4 %, p = 0.0035 Bonferroni corrected, Cohen's d = 0.59) and FFA (35.6 %, p = 0.024 Bonferroni corrected, Cohen's d = 0.60). We were not able to decode body identity across viewpoint from the body-responsive EBA (33.3 %, p = 0.53 uncorrected, Cohen's d = -0.012) or the face-responsive ATFA (33.4 %, p = 0.47 uncorrected, Cohen's d = 0.028). Using fMRI data from the viewpoint recognition task, we could decode body identity across viewpoint from the OFA (37.1 %, p < 0.0006 Bonferroni corrected, Cohen's d = 0.76), but not from any other ROI (EBA: 34.2 %, p = 0.16 uncorrected, Cohen's d = 0.22; FBA: 31.9 %, p = 0.96 uncorrected, Cohen's d = -0.30; FFA: 34.1 %, p = 0.21 uncorrected, Cohen's d = 0.21; ATFA: 33.3 %, p = 0.50 uncorrected, Cohen's d = 0.00).

We performed a whole-brain searchlight analysis to investigate if any other brain regions could decode body identity across viewpoint (Fig. 6B). From both the identity and the viewpoint task data, we could decode body identity across viewpoint from a large cluster in occipital cortex (including the early visual cortex). Using fMRI data from the identity recognition task, we could additionally decode body identity across viewpoint from the middle frontal gyrus, right anterior temporal cortex, right superior parietal cortex, right medial superior frontal gyrus, right insula cortex, right rolandic operculum and the left motor cortex (due to participants' button presses). Using fMRI data from the viewpoint recognition task, we could additionally decode body identity across viewpoint from the left fusiform gyrus, right superior parietal cortex, left caudate nucleus, left cingulum and left postcentral gyrus.

3.3.4.3. Viewpoint-invariant body identity classification: ID2 vs. ID3

As we found higher BOLD responses to ID1 as compared to ID2 and ID3 in the OFA, FFA and FBA during the identity response task (Fig. 3C), it is possible that our ability to decode body identity across viewpoint in these regions was the result of this higher BOLD

response to ID1. Therefore, we performed a follow-up analysis to test if these regions would be able to classify only body identities ID2 and ID3 across viewpoint (Fig. 5C). We performed the analysis using the same method as in *Section 3.3.4.2*, except that we only trained and tested the classifiers ability to distinguish between ID2 and ID3. We found we could decode body identity across viewpoint significantly above chance (50%) in the body-responsive FBA (52.6 %, p = 0.038 Bonferroni corrected, Cohen's d = 0.54) but not in the face-responsive OFA (51.1 %, p = 0.17 uncorrected, Cohen's d = 0.28) or FFA (51.5 %, 0.097 uncorrected, Cohen's d = 0.32) in this analysis.

3.3.5. Identity classification across face and body stimuli

Lastly, we performed multivoxel pattern analyses to investigate if any brain regions contain patterns of neural activity evoked by the three stimulus identities that could generalize across neural activity evoked by face and body stimuli (Fig. 7). To investigate this, we trained a linear SVM classifier to distinguish between patterns of neural activity evoked by the three face identities and then tested the classifier on its ability to decode the identity from neural activity evoked by the bodies (and vice-versa using the bodies for training the classifier and faces for testing it). As previously, we used a four-fold cross-validation method and performed the analysis twice separately using fMRI data recorded while participants performed the identity and viewpoint recognition tasks. We performed the analyses in face- and body-responsive ROIs (Fig. 7A) and in whole-brain searchlight analyses (Fig. 7B).

We were unable to decode identity across neural activity evoked by faces and bodies higher than chance-level (33 ⅓ %) in any of the ROIs we tested (Fig. 7A) using fMRI data from the identity recognition task (OFA: 34.2 %, p = 0.18 uncorrected, Cohen's d = 0.35; FFA: 34.0 %, p = 0.23 uncorrected, Cohen's d = 0.20; ATFA: 32.0 %, p = 0.93 uncorrected, Cohen's d = -0.40; EBA: 34.9 %, p = 0.042 uncorrected, Cohen's d = 0.58; FBA: 33.7 %, p = 0.34 uncorrected, Cohen's d = 0.16) or the viewpoint recognition task (OFA: 33.5 %, p = 0.43 uncorrected, Cohen's d = 0.051; FFA: 33.1 %, p = 0.58 uncorrected, Cohen's d = -0.071; ATFA: 33.4 %, p = 0.48 uncorrected, Cohen's d = 0.0086; EBA: 33.7 %, p = 0.33 uncorrected, Cohen's d = 0.11; FBA: 32.0 %, p = 0.93 uncorrected, Cohen's d = -0.67).

We performed whole-brain searchlight analyses to investigate if any other brain regions could decode identity across neural activity evoked by faces and bodies. Using fMRI

data from the identity recognition task (Fig. 7B) we could decode identity from the early visual cortex (MNI: 10, -94, 2), a region in the right inferior occipital cortex (MNI: 40, -84, -4) overlapping with the mean location of the OFA, the right parahippocampal cortex (MNI: 20, -4, -30) and a region in the right superior parietal cortex (MNI: 16, -56, 60). We could also decode identity from the left motor cortex due to participants' button presses. We were not able to decode identity from any regions using fMRI data from the viewpoint recognition task.

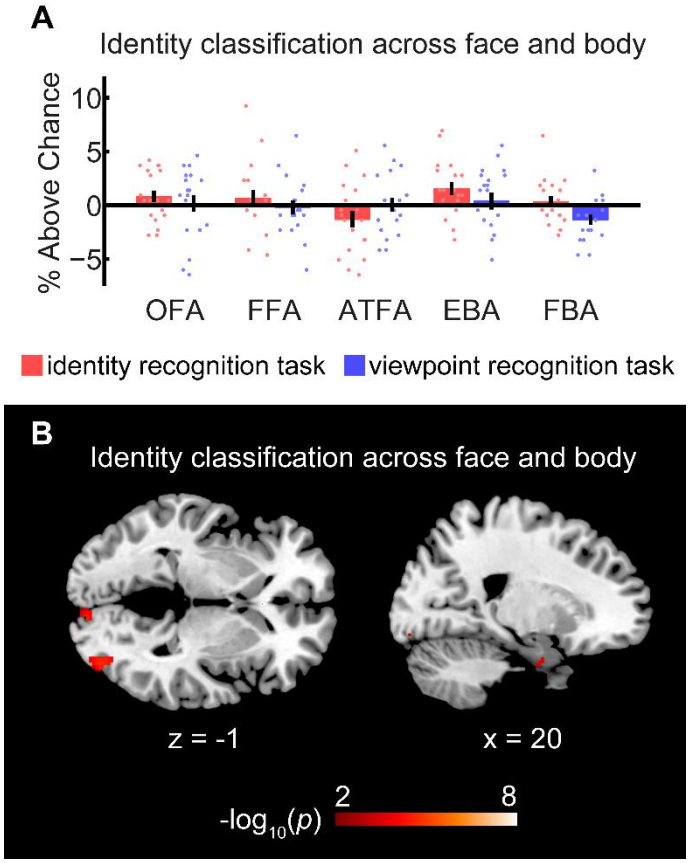

Figure 7. Classification of identity across neural activity evoked by faces and bodies. (A) shows classification of identity across face and body stimuli in face- and body-responsive ROIs. Scatter points show classification accuracies for individual participants and error bars show ±1 SEM. (B) shows classification of identity across face and body stimuli during the identity recognition task in a whole-brain searchlight analysis. The scale bar shows $-\log_{10}(p$ values) between 2 (p = 0.01) and 8 (p = 1 x10^{-8}), FDR corrected.

3.4. Discussion

In this study, we investigated the neural coding of face and body identity. Consistent with previous findings, we found that face identity could be decoded from neural activity in several distributed cortical regions (Anzellotti & Caramazza, 2015). We found that body identity could also be decoded from neural activity in several distributed cortical regions, and furthermore we found consistent decoding of body identity, including across viewpoint, from neural activity in the FBA, the right anterior temporal cortex and the middle frontal gyrus. We found we could decode identity in an abstract manner, across neural activity evoked by faces and bodies, from neural activity in the early visual cortex, right inferior occipital cortex, right parahippocampal cortex and right superior parietal cortex. These results provided new insights into how the brain encodes information about person identity.

3.4.1. Neural coding of face identity

We were able to classify face identity from the face-responsive ATFA and body-responsive EBA in our ROI analysis, and from regions in the early visual cortex, inferior occipital cortex, fusiform gyrus, superior parietal cortex, superior temporal cortex, parahippocampal cortex, right middle frontal gyrus, right anterior cingulum, right medial superior frontal gyrus and left inferior frontal gyrus in our searchlight analysis. Interestingly, we could decode face identity from these regions only when participants attended to the identity of the stimuli, not when they attended to the stimulus viewpoint. This suggests that face identity decoding in these regions was not based solely on visual features, as these were identical in both tasks (see Section 3.4.4 for a discussion of the behavioural task differences). Our face identity decoding results show consistence with previous findings demonstrating that face identity can be decoded from a number of distributed brain regions, including the ATFA, FFA, OFA, superior intraparietal sulcus and right inferior frontal cortex (Anzellotti & Caramazza, 2015; Anzellotti et al., 2014; Axelrod & Yovel, 2015; Goesaert & Op de Beeck, 2013; Guntupalli et al., 2016; Jeong & Xu, 2016; Kriegeskorte et al., 2007; Natu et al., 2010; Nestor et al., 2011).

We were unable to decode face identity across viewpoint from any brain region in this study. Electrophysiological recordings in macaque monkeys have shown that neurons in the anterior face patches respond to face identity across viewpoint (Freiwald & Tsao, 2010)

and human neuroimaging studies have shown that face identity can be decoded across viewpoint from human face-responsive regions (Anzellotti et al., 2014; Guntupalli et al., 2016). However, previous work has demonstrated that in some cases fMRI MVPA can fail to decode identity, even when electrophysiological recordings show that viewpoint-invariant identity information is present in the underlying neurons (Dubois, de Berker, & Tsao, 2015).

3.4.2. Neural coding of body identity

We could decode body identity from the body-responsive FBA and the face-responsive OFA and FFA in our ROI analyses, and from regions in the occipital cortex, fusiform gyrus, right anterior temporal cortex, superior parietal cortex, supramarginal gyrus, cingulum, precentral gyrus, caudate nucleus, inferior frontal gyrus, middle frontal gyrus, right insula cortex and right superior frontal gyrus in our searchlight analyses. Several of these regions have previously been shown to have higher responses when participants view the bodies of familiar people as compared to unfamiliar people (Hodzic et al., 2009). We could decode body identity from the OFA and occipital cortex regardless of the recognition task, suggesting this decoding could be based on differences in visual features between the different body identities. In contrast, we could only decode body identity from the right anterior temporal cortex, inferior frontal gyrus, middle frontal gyrus and right insula cortex when participants attended to identity, suggesting body identity responses in these regions are not driven purely by visual features, and that these responses were enhanced by participants' attention to identity (see Section 3.4.4).

We further investigated which brain regions encode body identity in a viewpoint-invariant manner. We could decode body identity across viewpoint from neural activity in the FBA, OFA and FFA in our ROI analyses, and from neural activity in the occipital cortex, middle frontal gyrus, right anterior temporal cortex, right superior parietal cortex, right medial superior frontal gyrus, right insula cortex, right rolandic operculum, left caudate nucleus, left cingulum and left postcentral gyrus in our searchlight analyses. We could decode body identity from the OFA and occipital cortex using fMRI data from both recognition tasks, suggesting that this decoding could be driven by visual features that are visible across different viewpoints. In contrast, we could decode body identity across viewpoint from the FBA, FFA, right anterior temporal cortex, middle frontal gyrus, right medial superior frontal gyrus, right insula cortex and right rolandic operculum only when

participants attended to identity, suggesting this decoding is driven by more abstract factors than visual features. However, we also found differences in the univariate responses to the three body identities in the FBA, FFA, right medial superior frontal gyrus and right insula cortex, which could have driven our significant decoding results in these regions. We did not find differences in the univariate responses to the three body identities in the right anterior temporal cortex, middle frontal gyrus or right rolandic operculum, suggesting that these responses were driven by differences in the pattern of neural responses only. Furthermore, in a follow-up analysis, we found that we could decode body identity across viewpoint in the FBA between two identities that showed no difference in their univariate responses in the FBA. Thus, altogether we find strongest evidence for neural encoding of body identity in a viewpoint-invariant manner in the FBA, the right anterior temporal cortex and the middle frontal gyrus, as these regions show consistent decoding of body identity in two analyses (body identity decoding and viewpoint-invariant body identity decoding), and these responses are not driven by low-level visual features or differences in univariate responses.

Several previous studies have found viewpoint-invariant face identity responses in the anterior temporal cortex (Anzellotti et al., 2014; Freiwald & Tsao, 2010; Guntupalli et al., 2016). Similarly, classification of body identity across viewpoint and pose was higher in a more anterior body patch in macaque temporal cortex than a more posterior body patch (Kumar et al., 2017), suggesting disentangling identity from viewpoint may be a general function performed by more anterior temporal regions. Our results showing that body identity can be decoded across viewpoint in the FBA and right anterior temporal cortex are consistent with a homology between the FBA and macaque anterior superior temporal sulcus body patch, and suggest that disentangling identity from viewpoint is also a general function of the human anterior temporal cortex. Furthermore, our results show a functional dissociation between the FBA and the EBA, as we were unable to decode body identity in the EBA in any of our analyses. Although a previous study showed body identity responses in both EBA and FBA using a repetition-suppression paradigm, the authors also showed evidence that feedback connectivity from the FBA to the EBA may have driven the body identity repetition suppression in the EBA (Ewbank et al., 2011). This is consistent with our finding of body identity encoding in the FBA and not the EBA.

3.4.3. Neural coding of identity across face and body

We could decode identity across neural activity evoked by face and body stimuli in the early visual cortex, the right inferior occipital cortex, the right parahippocampal cortex and the right superior parietal cortex. This abstract identity decoding was possible when participants attended to the stimulus identity, but not when they attended to viewpoint, showing that this decoding was not solely based on visual features and was enhanced by participants' attention to identity. Our ability to decode identity across the face and body in the early visual cortex suggests there may be feedback of identity information to the early visual cortex. Several previous studies have demonstrated there can be such feedback of high-level visual information to the early visual cortex (Bannert & Bartels, 2013; Grassi, Zaretskaya, & Bartels, 2017; Smith & Muckli, 2010; Williams et al., 2008; Zaretskaya, Anstis, & Bartels, 2013). Although we could not decode identity from our OFA ROI, we could decode identity from a region in the right inferior occipital cortex overlapping with the mean location of the right OFA. A previous study found viewpoint-invariant face identity responses in the OFA (Anzellotti et al., 2014), which, in combination with our results suggests the OFA contains some abstract identity encoding. We could also decode identity in an abstract manner in the right parahippocampal cortex. This region is known to be involved in memory and recollection (Eichenbaum, Yonelinas, & Ranganath, 2007), and has previously shown to be activated by recollection of contextual associations of faces and names (Kirwan & Stark, 2004). Our results suggest that this region also integrates identity information from the face and body. Finally, we could also decode identity in an abstract manner in the right superior parietal cortex. Consistent with this finding, previous work has identified abstract identity coding of faces and cars in the parietal cortex (Jeong & Xu, 2016). In combination, these results show consistence with previous brain regions found to respond to identity in an abstract manner and provide new insights into where in the brain identity information from the face and the body is combined.

3.4.4. Effect of recognition task

We could decode face identity and decode identity across face and body stimuli when participants attended to identity, but not when they attended to viewpoint. Similarly, we could decode body identity from the FBA, right anterior temporal cortex and middle frontal gyrus when participants attended to identity, but not when they attended to

viewpoint. Previous studies have demonstrated that attention to face identity enhances neural face identity responses (Dobs, Schultz, Bülthoff, & Gardner, 2018; Gratton, Sreenivasan, Silver, & D'Esposito, 2013). Furthermore, several studies reporting successful decoding of face identity used tasks where participants attended to identity (Anzellotti & Caramazza, 2015; Anzellotti et al., 2014; Guntupalli et al., 2016; Jeong & Xu, 2016; Nestor et al., 2011), whereas some studies reporting unsuccessful face identity decoding used tasks unrelated to face recognition (Dubois et al., 2015; Ramírez, Cichy, Allefeld, & Haynes, 2014). Our results, in combination with these previous studies, suggest that neural representations of face and body identity are enhanced by attention to identity, perhaps due to activation of identity-responsive neurons, and that this enhancement may be necessary to be able to decode identity from fMRI data.

3.5. Conclusion

We show, for the first time to our knowledge, that body identity can be decoded across viewpoint from neural activity in the body-responsive FBA, the right anterior temporal cortex and the middle frontal gyrus using MVPA. This result provides evidence that viewpoint-invariant identity coding may be a general function of more anterior regions of the human temporal cortex. Furthermore, we show that identity can be decoded in an abstract manner across neural activity evoked by faces and bodies in several regions previously associated with abstract identity coding. This provides new insights into the neural substrates encoding person identity information.

Acknowledgements

This research was supported by the Max Planck Society, Germany.

References

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), 905–913. https://doi.org/10.1016/j.neuroimage.2004.07.060

Anzellotti, S., & Caramazza, A. (2015). From Parts to Identity: Invariance and Sensitivity of Face Representations to Different Face Halves. *Cerebral Cortex*, *26*(5), 1900–1909. https://doi.org/10.1093/cercor/bhu337

Anzellotti, S., Fairhall, S. L., & Caramazza, A. (2014). Decoding Representations of Face Identity That are Tolerant to Rotation. *Cerebral Cortex*, *24*(8), 1988–1995. https://doi.org/10.1093/cercor/bht046

Axelrod, V., & Yovel, G. (2015). Successful Decoding of Famous Faces in the Fusiform Face Area. *Plos One*, *10*(2), 1–20. https://doi.org/10.1371/journal.pone.0117126

Bannert, M. M., & Bartels, A. (2013). Decoding the yellow of a gray banana. *Current Biology*, *23*(22), 2268–2272. https://doi.org/10.1016/j.cub.2013.09.016

Barton, J. J. S. (2008). Structure and function in acquired prosopagnosia: lessons from a series of 10 patients with brain damage. *Journal of Neuropsychology*, *2*(1), 197–225. https://doi.org/10.1348/174866407X214172

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Brooks, J. L. (2012). Counterbalancing for serial order carryover effects in experimental condition orders. *Psychological Methods*, *17*(4), 600–614. https://doi.org/10.1037/a0029310

Busigny, T., Van Belle, G., Jemel, B., Hosein, A., Joubert, S., & Rossion, B. (2014). Face-specific impairment in holistic perception following focal lesion of the right anterior temporal lobe. *Neuropsychologia*, *56*, 312–333. https://doi.org/10.1016/j.neuropsychologia.2014.01.018

Dobs, K., Schultz, J., Bülthoff, I., & Gardner, J. L. (2018). Task-dependent enhancement of facial expression and identity representations in human cortex. *NeuroImage*, *172*, 689–702. https://doi.org/10.1016/j.neuroimage.2018.02.013

Dubois, J., de Berker, A. O., & Tsao, D. Y. (2015). Single-Unit Recordings in the Macaque Face Patch System Reveal Limitations of fMRI MVPA. *The Journal of Neuroscience*, *35*(6), 2791–2802. https://doi.org/10.1523/JNEUROSCI.4037-14.2015

Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The Medial Temporal Lobe and Recognition Memory. *Annual Review of Neuroscience*, *30*, 123–152. https://doi.org/10.1146/annurev.neuro.30.051606.094328

Ewbank, M. P., Lawson, R. P., Henson, R. N., Rowe, J. B., Passamonti, L., & Calder, A. J. (2011). Changes in "Top-Down" Connectivity Underlie Repetition Suppression in the Ventral Visual Pathway. *The Journal of Neuroscience*, *31*(15), 5635–5642. https://doi.org/10.1523/JNEUROSCI.5013-10.2011

Fisher, C., & Freiwald, W. A. (2015). Whole-agent selectivity within the macaque face-processing system. *Proceedings of the National Academy of Sciences*, *112*(47), 14717–14722. https://doi.org/10.1073/pnas.1512378112

Foster, C., Zhao, M., Romero, J., Black, M. J., Mohler, B. J., Bartels, A., & Bülthoff, I. (2019). Decoding subcategories of human bodies from both body- and face-responsive cortical regions. *NeuroImage*, *202*, 1–13. https://doi.org/10.1016/j.neuroimage.2019.116085

Freiwald, W. A., & Tsao, D. Y. (2010). Functional Compartmentalization and Viewpoint Generalization Within the Macaque Face-Processing System. *Science*, *330*(6005), 845–851. https://doi.org/10.1126/science.1194908

Gauthier, I., Tarr, M. J., Moylan, J., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). The fusiform "face

area" is part of a network that processes faces at the individual level. *Journal of Cognitive Neuroscience*, *12*(3), 495–504. https://doi.org/10.1162/089892900562165

Goesaert, E., & Op de Beeck, H. P. (2013). Representations of facial identity information in the ventral visual stream investigated with multivoxel pattern analyses. *The Journal of Neuroscience*, *33*(19), 8549–8558. https://doi.org/10.1523/JNEUROSCI.1829-12.2013

Grassi, P. R., Zaretskaya, N., & Bartels, A. (2017). Scene segmentation in early visual cortex during suppression of ventral stream regions. *NeuroImage*, *146*, 71–80. https://doi.org/10.1016/j.neuroimage.2016.11.024

Gratton, C., Sreenivasan, K. K., Silver, M. A., & D'Esposito, M. (2013). Attention selectively modifies the representation of individual faces in the human brain. *The Journal of Neuroscience*, *33*(16), 6979–6989. https://doi.org/10.1523/JNEUROSCI.4142-12.2013

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. https://doi.org/10.1038/nn1224

Guntupalli, J. S., Wheeler, K. G., & Gobbini, M. I. (2016). Disentangling the Representation of Identity from Head View Along the Human Face Processing Pathway. *Cerebral Cortex*, *27*(1), 46–53. https://doi.org/10.1093/cercor/bhw344

Hadjikhani, N., & de Gelder, B. (2002). Neural basis of prosopagnosia: An fMRI study. *Human Brain Mapping*, *16*(3), 176–182. https://doi.org/10.1002/hbm.10043

Hahn, C. A., O'Toole, A. J., & Phillips, P. J. (2015). Dissecting the time course of person recognition in natural viewing environments. *British Journal of Psychology*, *107*(1), 117–134. https://doi.org/10.1111/bjop.12125

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*(6), 223–233. https://doi.org/10.1016/S1364-6613(00)01482-0

Hebart, M. N., Görgen, K., & Haynes, J.-D. (2015). The Decoding Toolbox (TDT): a versatile software package for multivariate analyses of functional imaging data. *Frontiers in Neuroinformatics*, *8*, 1–18. https://doi.org/10.3389/fninf.2014.00088

Hodzic, A., Kaas, A., Muckli, L., Stirn, A., & Singer, W. (2009). Distinct cortical networks for the detection and identification of human body. *NeuroImage*, *45*(4), 1264–1271. https://doi.org/10.1016/j.neuroimage.2009.01.027

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), 80–84. https://doi.org/10.1038/71152

Jeong, S. K., & Xu, Y. (2016). Behaviorally Relevant Abstract Object Identity Representation in the Human Parietal Cortex. *The Journal of Neuroscience*, *36*(5), 1607–1619. https://doi.org/10.1523/JNEUROSCI.1016-15.2016

Jonas, J., Rossion, B., Brissart, H., Frismand, S., Jacques, C., Hossu, G., … Maillard, L. (2015). Beyond the core face-processing network: Intracerebral stimulation of a face-selective area in the right anterior fusiform gyrus elicits transient prosopagnosia. *Cortex*, *72*, 140–155. https://doi.org/10.1016/j.cortex.2015.05.026

Kirwan, C. B., & Stark, C. E. L. (2004). Medial temporal lobe activation during encoding and retrieval of novel face-name pairs. *Hippocampus*, *14*(7), 919–930. https://doi.org/10.1002/hipo.20014

Kleiner, M., Brainard, D., & Pelli, D. (2007). "What's new in Psychtoolbox-3?" In *Perception 36 ECVP Abstract Supplement.*

Kriegeskorte, N., Formisano, E., Sorger, B., & Goebel, R. (2007). Individual faces elicit distinct response patterns in human anterior temporal cortex. *Proceedings of the National Academy of Sciences*, *104*(51), 20600–20605. https://doi.org/10.1073/pnas.0705654104

Kumar, S., Popivanov, I. D., & Vogels, R. (2017). Transformation of Visual Representations Across Ventral

Stream Body-selective Patches. *Cerebral Cortex*, *29*(1), 215–229. https://doi.org/10.1093/cercor/bhx320

Li, T., Bolkart, T., Black, M. J., Li, H., & Romero, J. (2017). Learning a model of facial shape and expression from 4D scans. *ACM Transactions on Graphics*, *36*(6), 1–17. https://doi.org/10.1145/3130800.3130813

Loffler, G., Yourganov, G., Wilkinson, F., & Wilson, H. R. (2005). fMRI evidence for the neural representation of faces. *Nature Neuroscience*, *8*(10), 1386–1390. https://doi.org/10.1038/nn1538

Loper, M., Mahmood, N., Romero, J., Pons-Moll, G., & Black, M. J. (2015). SMPL : A Skinned Multi-Person Linear Model. *ACM Trans. Graphics (Proc. SIGGRAPH Asia)*, *34*(6), 248:1--248:16. https://doi.org/10.1145/2816795.2818013

Nasr, S., & Tootell, R. B. H. (2012). Role of fusiform and anterior temporal cortical areas in facial recognition. *NeuroImage*, *63*(3), 1743–1753. https://doi.org/10.1016/j.neuroimage.2012.08.031

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. https://doi.org/10.1162/jocn.2009.21312

Nestor, A., Plaut, D. C., & Behrmann, M. (2011). Unraveling the distributed neural code of facial identity through spatiotemporal pattern analysis. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(24), 9998–10003. https://doi.org/10.1073/pnas.1102433108

O'Toole, A. J., Phillips, P. J., Weimer, S., Roark, D. A., Ayyad, J., Barwick, R., & Dunlop, J. (2011). Recognizing people from dynamic and static faces and bodies: Dissecting identity with a fusion approach. *Vision Research*, *51*(1), 74–83. https://doi.org/10.1016/j.visres.2010.09.035

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442. https://doi.org/10.1163/156856897X00366

Pitcher, D., Charles, L., Devlin, J. T., Walsh, V., & Duchaine, B. (2009). Triple Dissociation of Faces, Bodies, and Objects in Extrastriate Cortex. *Current Biology*, *19*(4), 319–324. https://doi.org/10.1016/j.cub.2009.01.007

Premereur, E., Taubert, J., Janssen, P., Vogels, R., & Vanduffel, W. (2016). Effective Connectivity Reveals Largely Independent Parallel Networks of Face and Body Patches. *Current Biology*, *26*(24), 3269–3279. https://doi.org/10.1016/j.cub.2016.09.059

Ramírez, F. M., Cichy, R. M., Allefeld, C., & Haynes, J.-D. (2014). The neural code for face orientation in the human fusiform face area. *The Journal of Neuroscience*, *34*(36), 12155–12167. https://doi.org/10.1523/JNEUROSCI.3156-13.2014

Rice, A., Phillips, P. J., Natu, V., An, X., & O'Toole, A. J. (2013). Unaware Person Recognition From the Body When Face Identification Fails. *Psychological Science*, *24*(11), 2235–2243. https://doi.org/10.1177/0956797613492986

Rice, A., Phillips, P. J., & O'Toole, A. (2013). The role of the face and body in unfamiliar person identification. *Applied Cognitive Psychology*, *27*(6), 761–768. https://doi.org/10.1002/acp.2969

Robbins, R. A., & Coltheart, M. (2012). The effects of inversion and familiarity on face versus body cues to person recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(5), 1098–1104. https://doi.org/10.1037/a0028584

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. https://doi.org/10.1038/nn1370

Schwarzlose, R. F., Baker, C. I., & Kanwisher, N. (2005). Separate Face and Body Selectivity on the Fusiform Gyrus. *The Journal of Neuroscience*, *25*(47), 11055–11059. https://doi.org/10.1523/JNEUROSCI.2621-05.2005

Smith, F. W., & Muckli, L. (2010). Nonstimulated early visual areas carry information about surrounding context. *Proceedings of the National Academy of Sciences*, *107*(46), 20099–20103.

https://doi.org/10.1073/pnas.1000233107

Williams, M. A., Baker, C. I., Op De Beeck, H. P., Shim, W. M., Dang, S., Triantafyllou, C., & Kanwisher, N. (2008). Feedback of visual object information to foveal retinotopic cortex. *Nature Neuroscience*, *11*(12), 1439–1445. https://doi.org/10.1038/nn.2218

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), 1830–1839. https://doi.org/10.1152/jn.00155.2004

Zaretskaya, N., Anstis, S., & Bartels, A. (2013). Parietal Cortex Mediates Conscious Perception of Illusory Gestalt. *The Journal of Neuroscience*, *33*(2), 523–531. https://doi.org/10.1523/JNEUROSCI.2905-12.2013

4. An abstract neural code in occipital cortex for person orientation

Abstract

We can easily recognise the orientation of people we see, for example whether their body is facing us or turned toward another direction. This ability is important for us to understand other people's actions, social interactions and intentions. Previous neuroimaging studies have shown that several brain regions in occipitotemporal cortex have different patterns of response when we view faces or bodies of different orientations. However, it is often unclear from these studies whether these different patterns of response are driven by an abstract encoding of orientation or simply by low-level visual features that correlate with different face and body orientations. Furthermore, no study so far has directly compared neural responses to face and body orientations in the human brain. In the present study, we recorded participants' brain activity using fMRI while they viewed faces and bodies from three different orientations. This allowed us to compare which brain regions process face and body orientation and, as faces and bodies vary considerably in their low-level visual properties, investigate if any regions encode person orientation in an abstract manner. We found that that the occipital face area (OFA) and extrastriate body area (EBA) respond to person orientation in an abstract code that generalized across neural activity evoked by faces and bodies. Furthermore, we found that the fusiform face area (FFA) and fusiform body area (FBA) responded to face orientation but not body orientation, suggesting that orientation responses in the fusiform gyrus are face-specific. Our results show that early face- and body-responsive regions encode person orientation in a manner abstracted from low-level visual features, suggesting that orientation processing is an important function of these occipital regions.

Keywords: orientation, viewpoint, body recognition, face recognition, fMRI

4.1. Introduction

The ability to process the orientation of people we see is important for us to be able to understand how other people interact with us and the world around them. For example, we know that if a person is oriented towards an object they may intend to interact with it or if they are oriented towards us, the viewer, they may want to socially interact with us. Humans use information from the face and body to determine a person's orientation, and in addition use information from their eye gaze to determine their direction of attention. Psychological research has shown there are interactions between face and body orientation information (Moors, Germeys, Pomianowska, & Verfaillie, 2015) and face orientation and gaze direction information (Gibson & Pick, 1963; Wollaston, 1824).

Neuroimaging and electrophysiology studies have investigated which brain regions process information about face and body orientation in humans and macaque monkeys. For face orientation, human neuroimaging studies have found different patterns of response to different face orientations in the face-responsive occipital face area (OFA), fusiform face area (FFA) and posterior superior temporal sulcus (pSTS), object-responsive lateral occipital (LO) area and early visual cortex (Axelrod & Yovel, 2012; Guntupalli, Wheeler, & Gobbini, 2016; Kietzmann, Swisher, König, & Tong, 2012; Natu et al., 2010; Ramírez, Cichy, Allefeld, & Haynes, 2014). Similarly, for macaque monkeys, neurons in the posterior face-responsive patches and the anterior superior temporal sulcus (aSTS) have been shown to respond to specific face orientations (Dubois, de Berker, & Tsao, 2015; Freiwald & Tsao, 2010; Perrett et al., 1985; Wachsmuth, Oram, & Perrett, 1994). For body orientation, human neuroimaging studies have shown that the body-responsive extrastriate body area (EBA) and fusiform body area (FBA) are sensitive to body orientation (Chan, Peelen, & Downing, 2004; Ewbank et al., 2011; Taylor, Wiggett, & Downing, 2010). In macaques, body orientation can be decoded from both the middle and anterior superior temporal sulcus body patches (Kumar, Popivanov, & Vogels, 2017) and body orientation responsive neurons have been identified in the macaque aSTS (Wachsmuth et al., 1994).

These studies demonstrate that many occipitotemporal regions respond to face and/or body orientation. However, for many of these studies it is unclear whether the neural responses reflect an abstract high-level encoding of orientation direction or

responses to low-level visual features that correlate between face or body images of different individuals with the same orientation. One study found that some neurons in the macaque aSTS respond to orientation direction in an abstract manner (Wachsmuth et al., 1994). These neurons responded to a particular orientation direction when it was shown from both face and body images shown separately, which contain very different low-level visual features. However, as the authors only recorded from the aSTS it is not clear if any other brain regions would encode orientation in an abstract manner. Furthermore, as most studies have investigated orientation responses separately for faces and bodies it is not possible to directly compare which brain regions are involved in orientation processing of both faces and bodies.

In this study, we investigated the neural responses to different face and body orientations in face- and body-responsive brain regions and across the whole brain in searchlight analyses. We recorded participants' brain activity using fMRI while they viewed images of faces and bodies of three people from three different orientations. We then trained linear support vector machine (SVM) classifiers to distinguish between patterns of neural activity evoked by the three stimulus orientations and used them to predict stimulus orientation in a separate set of test data. Firstly, we performed these analyses separately for neural activity evoked by faces and bodies in order to directly compare which brain regions respond to face and body orientation. Secondly, we trained classifiers on neural activity evoked by faces and tested them on neural activity evoked by bodies, and vice-versa, in order to investigate which brain regions contain abstract responses to orientation that can generalize across neural activity evoked by faces and bodies. Thirdly, we also varied participants' attention during the experiment. In half of the dataset participants were instructed to respond to the orientation of the stimuli, and in the other half of the dataset the identity of the stimuli. This allowed us to compare if there would be differences in participants' neural responses depending on whether they attended to the stimulus orientation or identity.

4.2. Materials and methods

Data analyses presented here were conducted using a dataset that was collected as part of a larger study. Here we present novel findings from behavioural and fMRI data analyses investigating behavioural and neural responses to face and body orientation.

4.2.1. Participants

20 participants (14 female, 21-51 years old) completed the experiment. All participants provided written informed consent prior to the experiment, and the experimental procedure was approved by the local ethics committee of the University Clinic Tübingen.

4.2.2. Stimuli

4.2.2.1. Main experiment stimuli

The experimental stimuli were images of faces and bodies shown from three different orientations, 0° (front), 45° and 90° (profile). Examples of the stimuli are shown in Fig. 1A. Face and body stimuli were created using face and body scans of three female individuals that were registered to a 3D facial shape and expression model for the face stimuli (Li, Bolkart, Black, Li, & Romero, 2017) and a 3D body shape and pose model for the body stimuli (Loper, Mahmood, Romero, Pons-moll, & Black, 2015). Body stimuli were shown in a standard A-pose (see Fig. 1A) and faces had a neutral expression. Stimuli were shown in colour, and for body stimuli a grey rectangle was placed over the face, in order to exclude any face information from the body images. During the experiment face and body stimuli were shown at three different image sizes. Face stimuli had mean widths and heights of 4.4° x 6.4°, 3.6° x 5.2° and 2.8° x 4.0° of visual angle, and body stimuli had mean widths and heights of 3.2° x 7.7°, 2.6° x 6.2° and 2.0° x 4.8° of visual angle.

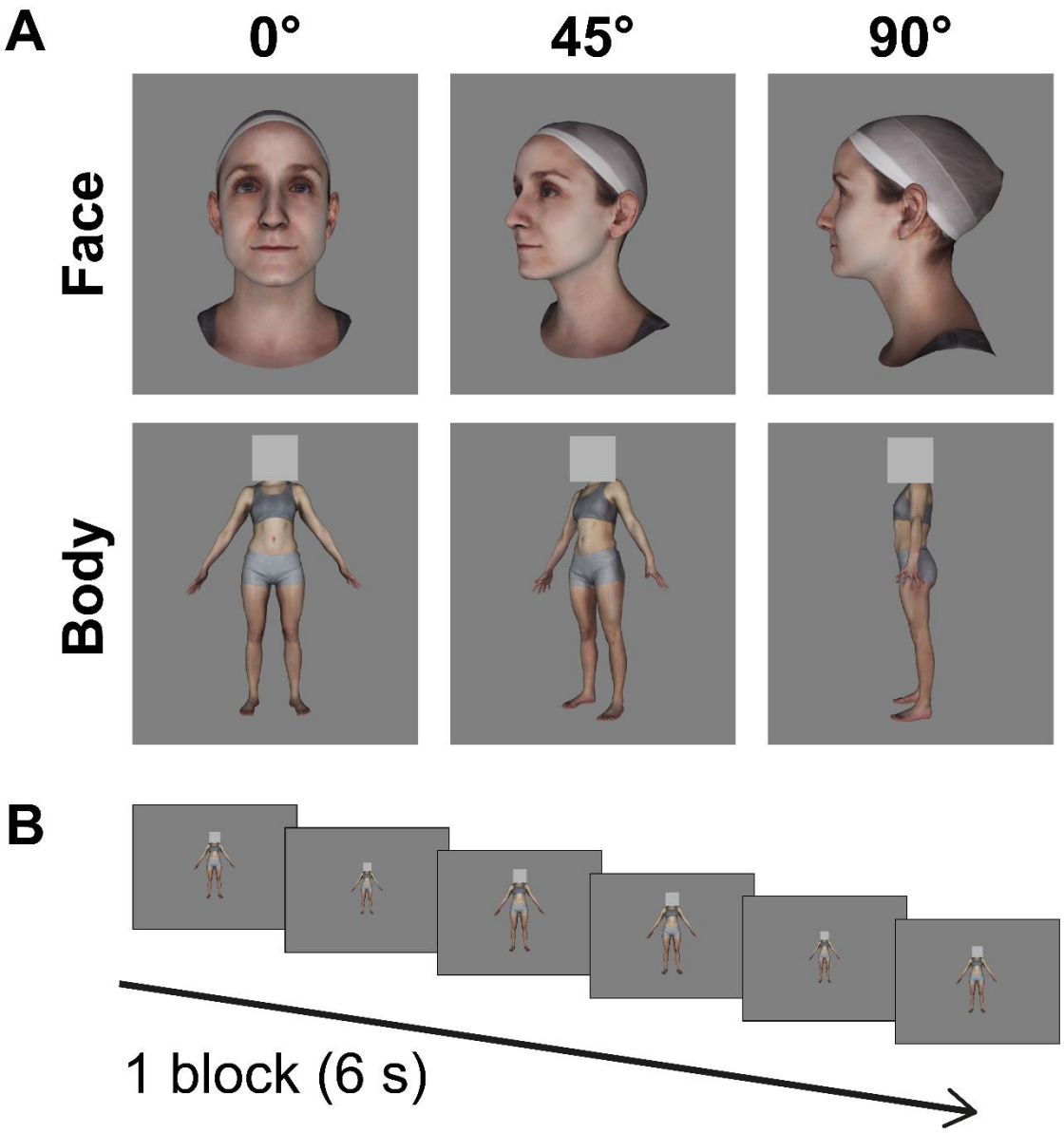

Figure 1. Experimental stimuli and example stimulus block. (A) Face and body images were shown from three different orientations: 0° (front), 45° and 90° (profile). (B) Stimuli were shown in a block design during the experiment, where stimuli within a block were all from one condition (i.e. face or body, one orientation, one identity) and varied in their image-size (three different image-sizes, each shown twice, presented in a random order). Each block was followed by 2 s fixation.

4.2.2.2. Localizer stimuli

Stimuli for the localizer experiment consisted of grayscale images of faces, headless bodies, objects and phase-scrambled images. Phase-scrambled images were Fourier-scrambled versions of a collage image containing the face and headless body images.

4.2.3. Experimental design

Participants lay supine in the MRI scanner and viewed the stimuli on a screen positioned 92 cm behind their head, which they viewed via a mirror attached to the head coil. The stimuli were presented using a projector (resolution 1920x1080), and the screen spanned 25° x 14° of visual angle in horizontal and vertical directions respectively. The experiment was programmed on Ubuntu 17.10 with Matlab 2017a using the Psychophysics Toolbox extensions (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007).

4.2.3.1. Main experiment procedure

Each participant completed eight fMRI runs, where each run contained 18 conditions of a 2 (face or body) x 3 (orientation) x 3 (identity) factorial design. Stimuli were presented in a block design (Fig. 1B) where each block contained 6 images that were all from the same condition (i.e. face or body, one orientation, one identity). Images within a block were shown at three different image-sizes with scale factors of 1, 1.3 and 1.6 (i.e. the largest image-size was 1.6 times both the width and height of the smallest image-size). Each block contained 2 repetitions of each image-size, shown in a random order. Each image in the block was shown for 0.9 s and a 0.1 s blank grey screen followed each image. Following each block was a 2 s fixation before the next block began. Each run contained 54 blocks (3 repetitions per condition), where 18 blocks (one per condition) were presented in a random order preceded by and followed by 8 s of fixation.

4.2.3.2. Main experiment task

In half of the fMRI runs participants were instructed to respond at the end of each block to which orientation (i.e. 0°, 45° or 90°) was presented in the block. In the other half of the runs participants were instructed to respond at the end of the block to which identity was presented in the block. Participants were trained to recognise the three identities prior

to the fMRI experiment. Participants pressed a button with one of three fingers to indicate the orientation or identity shown in the block.

In all runs participants performed an additional attention task to keep their attention on the stimuli throughout the block. Participants were instructed to respond by pressing a button with their thumb immediately whenever they saw an image of the smallest of the three image-sizes.

4.2.3.3. fMRI localizer experiment procedure

Participants completed one run of a localizer experiment following the main experiment. Data from this localizer was used to define face- and body-responsive regions of interest. The localizer consisted of four conditions (faces, bodies, objects and phase-scrambled images) that were presented in a block design. Faces, bodies and objects were shown in front of the phase-scrambled images so that the size of the visual-field stimulation was the same for every image. Each block contained 8 images from one condition where each image was presented for 1.8 s and was followed by a 0.2 s blank grey screen. Conditions were presented in a carryover counterbalanced sequence, so that each condition was preceded by each condition an equal number of times (Brooks, 2012).

During the localizer, participants performed a one-back matching task on the images, to ensure they kept their attention on the stimuli. Image repetitions occurred once every 9 seconds on average.

4.2.4. Imaging parameters

Images were acquired using a 3T Siemens Prisma scanner with a 64-channel head coil (Siemens, Erlangen, Germany). Functional T2* echoplanar images (EPI) were acquired using a sequence with the following parameters; multiband acceleration factor 2, GRAPPA acceleration factor 2, TR 1.84 s, TE 30 ms, flip angle 79°, FOV 192x192 mm. Volumes consisted of 60 slices, with an isotropic voxel size of 2x2x2 mm. The first 8 volumes of each run were discarded to allow for equilibration of the T1 signal. For each participant a high-resolution T1-weighted anatomical scan was acquired with the following parameters; TR 2 s, TE 3.06 ms, FOV 232x256 mm, 192 slices, isotropic voxel size of 1x1x1 mm.

4.2.5. MRI data preprocessing

MRI data was preprocessed with SPM12 (http://www.fil.ion.ucl.ac.uk/spm/). All functional images were slice-time corrected, realigned and coregistered to the anatomical image. ROI and whole-brain searchlight analyses were conducted on the unsmoothed data in subject space. The resulting searchlight maps were normalised to MNI (Montreal Neurological Institute) space, and spatially smoothed with a 6 mm Gaussian kernel to allow for comparisons across participants. For the whole-brain univariate analyses the data was normalized to MNI space and spatially smoothed with a 6 mm Gaussian kernel. Localizer data was kept in subject-space and spatially-smoothed with a 6 mm Gaussian kernel.

4.2.6. Definition of regions of interest

We defined face- and body-responsive regions of interests (ROIs) using fMRI data from our localizer experiment. We defined four face-responsive ROIs, the occipital face area (OFA), the fusiform face area (FFA) the posterior superior temporal sulcus (pSTS) and the anterior temporal face area (ATFA) and two body-responsive ROIs, the extrastriate body area (EBA) and the fusiform body area (FBA). For each participant, we initially attempted to define face-responsive ROIs using the contrast faces > objects and body-responsive ROIs using the contrast bodies > objects. If we could not define the ROIs using these contrasts then we attempted to define them using the contrasts faces > scrambled images and bodies > scrambled images. We first attempted to define ROIs using a threshold of p <0.001 (uncorrected), and then reduced this threshold to $p < 0.01$ (uncorrected) if the ROI could not be defined using the first threshold.

Table 1

Average MNI coordinates and volume of each ROI, ± standard deviations. N indicates the number of participants each ROI was identified in. ROI analyses were conducted in subject space and then ROIs were subsequently normalised to generate the MNI coordinates.

ROI	hem	x	y	z	Volume (mm^3)	N
OFA	left	-35 ± 6.7	-85 ± 5.7	-11 ± 3.5	770 ± 379.9	20
	right	38 ± 4.0	-81 ± 5.8	-10 ± 3.3	1009 ± 378.6	20
FFA	left	-40 ± 2.8	-55 ± 6.1	-20 ± 2.9	771 ± 354.7	20
	right	43 ± 3.3	-52 ± 4.2	-18 ± 2.4	1073 ± 392.5	20
pSTS	left	-50 ± 6.4	-62 ± 8.8	17 ± 10.2	519 ± 541.2	19
	right	53 ± 5.5	-54 ± 10.0	12 ± 8.9	732 ± 403.1	20
ATFA	left	-34 ± 5.4	-12 ± 6.5	-33 ± 6.6	172 ± 117.4	15
	right	34 ± 5.8	-8 ± 5.5	-37 ± 5.8	335 ± 265.6	18
EBA	left	-44 ± 3.7	-78 ± 5.3	3 ± 6.5	900 ± 473.3	20
	right	49 ± 2.3	-70 ± 3.0	0 ± 4.6	1632 ± 503.2	20
FBA	left	-39 ± 4.2	-50 ± 6.3	-20 ± 2.9	680 ± 456.7	19
	right	41 ± 4.1	-50 ± 5.5	-18 ± 3.0	1105 ± 572.2	20

4.2.7. Behavioural analyses

Participants were instructed to respond with a button press at the end of each block to indicate which orientation or identity was presented in the block (one half of blocks orientation task, the other half of blocks identity task). We calculated our participants' behavioural performance using accuracy (% correct). To investigate if there were any differences in the detection of orientation or identity of stimuli from the three different orientations we performed one-way repeated-measures ANOVAs with three levels (0°, 45° and 90° stimulus orientation). We corrected for non-sphericity where necessary following a Mauchly's test of sphericity. Following any significant ANOVA results, we performed follow-

up paired *t*-tests between all combinations of the three orientation conditions to determine exactly which conditions showed differences in behavioural performance.

4.2.8. Multivoxel pattern analyses (MVPA)

We conducted multivoxel pattern analyses (MVPA) to investigate which brain regions contain different patterns of activity to faces and bodies of different orientations (0°, 45° and 90°). Following preprocessing, fMRI data was modelled with a General Linear Model (GLM) using SPM12, where the neural responses to each block were modelled as separate regressors in the GLM. MVPA analyses were then performed on the beta weight images with The Decoding Toolbox (Hebart, Görgen, & Haynes, 2015) using a linear support vector machine classifier (LIBSVM). We performed feature-scaling on the input data using z-score normalization and set any outlier values (values that were greater than 2 standard deviations from the mean) to 2 or -2. We estimated the mean and standard deviation for feature-scaling using the training data and then applied these values to the test data.

In a first set of analyses, we aimed to determine which brain regions contain separable patterns of activity to faces of different orientations and bodies of different orientations. Thus, we analysed fMRI data evoked by face and body stimuli separately. We analysed fMRI data from the two behavioural tasks separately, thus we used 4 runs of fMRI data per analysis. In each analysis, we trained a linear SVM classifier to distinguish between neural activity evoked by the three stimulus orientations. We trained the classifier using neural activity data evoked by 2 of the 3 stimulus identities and from 3 of the 4 runs of fMRI data. We then tested the classifier on its ability to predict the three stimulus orientations from neural activity data evoked by the third stimulus identity from the 4^{th} left out run of fMRI data. A brain region showing higher than chance decoding performance in this analysis would show that the region contains separable patterns of neural activity to the three different stimulus orientations, and that these patterns are invariant with respect to the identity of the stimulus. We used a 4-fold cross-validation procedure where we repeated the analysis 4 times with each run used once as the held out test dataset. We also repeated the analysis 3 times with each stimulus identity used as the held out test identity once. The final decoding accuracy was determined by averaging over the 4 cross-validation and 3 stimulus identity combinations.

In a second set of analyses, we aimed to determine which brain regions contain separable patterns of neural activity evoked by the three stimulus orientations that could generalize across neural activity evoked by face and body stimuli. We again analysed fMRI data from the two behavioural tasks separately, and thus used 4 runs of data per analysis. We trained a linear SVM classifier to distinguish between neural activity evoked by the three stimulus orientations, using neural activity data evoked by face stimuli from 3 of the 4 runs of fMRI data. We then tested the classifier on its ability to predict the three stimulus orientations from neural activity data evoked by body stimuli from the 4th left out run of fMRI data. We again used a 4-fold cross-validation procedure where we repeated the analysis 4 times with each run used once as the held out test dataset. In addition, we repeated the analysis but using neural activity data evoked by body stimuli as the training set and neural activity data evoked by face stimuli as the test dataset. The final decoding accuracy was determined by averaging over the 4 cross-validation and the two training and test dataset combinations.

We conducted all MVPA analyses in ROIs and whole-brain searchlight analyses. For ROI analyses we determined statistical significance using permutation testing. For each ROI we repeated each analysis 10,000 times with the condition labels assigned in a random order, in order to generate a null distribution of classification accuracies that would be expected by chance. We assessed significance by comparing how often in this null distribution we obtained a mean decoding performance equal to or greater than our actual mean decoding performance. We tested for significance using a threshold of $p < 0.05$, using a Bonferroni correction to adjust for multiple comparisons (N = 6 ROIs tested).

Whole-brain searchlight analyses were performed in subject-space using 4-voxel radius spheres, which were centred around each voxel in the brain once. Thus for each participant and each analysis we obtained a whole-brain map of classification accuracies. These maps were then normalised to MNI space and smoothed with a 6 mm Gaussian kernel to allow for comparisons across participants. We used SnPM13 (http://warwick.ac.uk/snpm) to assess significance using nonparametric permutation tests (Nichols & Holmes, 2001) with 10,000 permutations and 6 mm FWHM variance smoothing. We tested for significance using a one-sided *t*-test with a threshold of $p < 0.05$, family-wise error rate (FWE) corrected for multiple comparisons.

4.2.9. Univariate analyses

We conducted univariate analyses to investigate if there were brain regions that showed different mean overall levels of neural response to faces or bodies of different orientations. Following preprocessing, we modelled the fMRI data with a GLM using SPM12, where the neural responses to each condition were modelled as separate regressors. We performed univariate analyses in ROIs and in whole-brain analyses. We tested for differences in the level of neural activity to the three different stimulus orientations using one-way repeated measures ANOVAs with three levels (0°, 45° and 90° stimulus orientation). For ROI analyses we assessed significance using a threshold of $p < 0.05$, corrected for non-sphericity where necessary following a Mauchly's test of sphericity, and Bonferroni-corrected for multiple comparisons (N = 6). In ROIs showing significant differences in the ANOVA analyses, we performed follow up paired *t*-tests between the different stimulus orientations to determine which conditions showed differences in neural activation. For whole-brain analyses we assessed significance with a threshold of $p < 0.05$, FWE corrected.

4.3. Results

4.3.1. Behavioural results

We measured participants accuracy in detecting the orientation and identity of stimuli of the three different stimulus orientations (0°, 45° and 90°) during the fMRI experiment using accuracy (% correct). The results are shown in Figure 2.

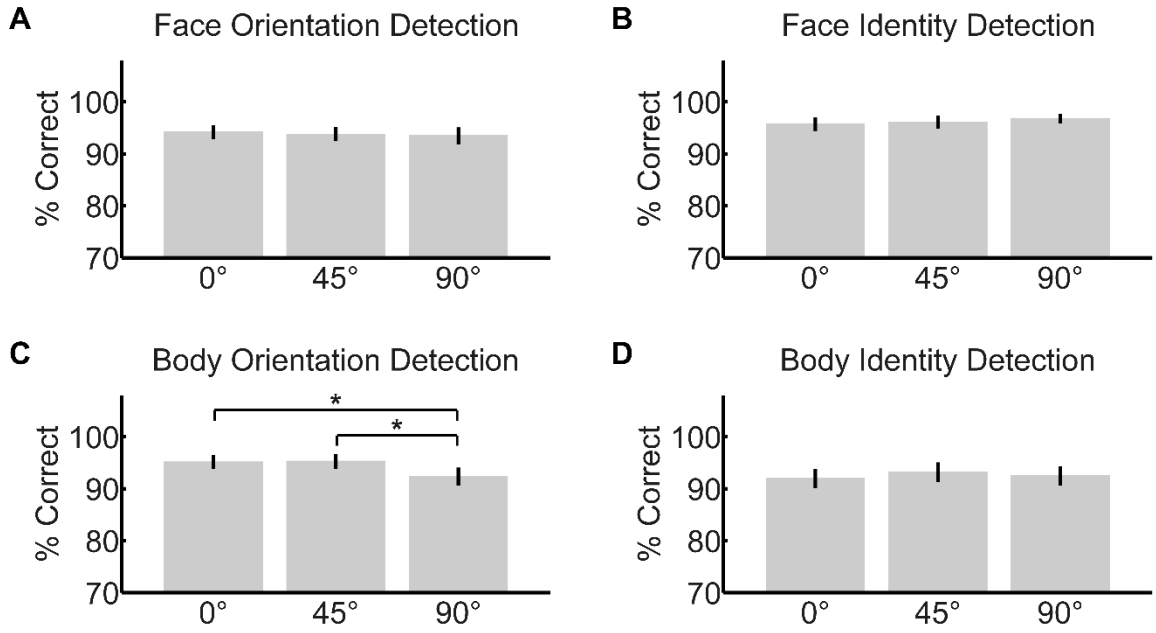

Figure 2. Accuracy (% correct) in detection of orientation and identity for stimuli from the three different orientation conditions. (A) and (C) show participants accuracy in detecting the orientation of face (A) and body (C) stimuli. (B) and (D) show participants accuracy in detecting the identity of face (B) and body (D) stimuli. Error bars indicate ±1 SEM. * indicates $p < 0.05$.

4.3.1.1. Orientation detection

Participants showed a high accuracy in detecting the orientation of both face and body stimuli across all conditions (face stimuli: 93.8 % correct; body stimuli: 94.3 % correct). A one-way repeated-measures ANOVA showed there were no differences in orientation detection performance between face stimuli of the three different orientations ($F_{2,38} = 0.23$, $p = 0.79$, $\eta_p^2 = 0.012$). In contrast, a one-way repeated-measures ANOVA showed there were

differences in orientation detection between body stimuli of the three different orientations ($F_{2,38}$ = 3.54, p = 0.039, η_p^2 = 0.16). Follow-up paired t-tests showed that detection accuracy of 90° bodies was lower than the detection accuracy of both 45° body stimuli (M = -2.9 %, SE = 1.31, t_{19} = -2.2, p = 0.039, Cohen's d = -0.50) and 0° body stimuli (M = -2.8 %, SE = 1.21, t_{19} = -2.3, p = 0.033, Cohen's d = -0.54), but there was no difference in detection accuracy between 45° and 0° body stimuli (M = 0.1 %, SE = 1.18, t_{19} = 0.12, p = 0.91, Cohen's d = 0.026). We note that detection accuracy was very high for all three body orientations (0°: 95.1 %, 45°: 95.3 %, 90°: 92.4 %), showing that participants could easily detect all three body orientations.

4.3.1.2. Identity detection

Participants showed high accuracy in the detection of identity for both face (96.2 %) and body (92.5 %) stimuli. One-way repeated-measures ANOVAs showed there were no differences in identity detection between stimuli of the three different orientations for both face stimuli ($F_{2,38}$ = 0.68, p = 0.51, η_p^2 = 0.035) and body stimuli ($F_{2,38}$ = 0.30, p = 0.74, η_p^2 = 0.016). This shows that participants could detect the stimulus identities equally well regardless of the orientation of the stimuli.

4.3.2. Neural responses to face orientation

4.3.2.1. Classification analyses

We first investigated which brain regions have different patterns of neural activity evoked by different face orientations that could generalize across face identity. We trained a linear SVM classifier to distinguish between patterns of neural activity evoked by the three face orientations of two identities. We then tested the classifier on its ability to decode the face orientation of a third identity, using neural activity data in a left out run of data. We used a leave one run out cross validation method, and also repeated the analysis with each identity used once as the test identity. We performed the analysis in face- and body-responsive ROIs as well as in searchlight analyses across the whole brain. We conducted the analysis twice separately for data where participants responded to the orientation of the stimuli (orientation task) and for data where they responded to the identity of the stimuli (identity task). The results are shown in Fig. 3A-D.

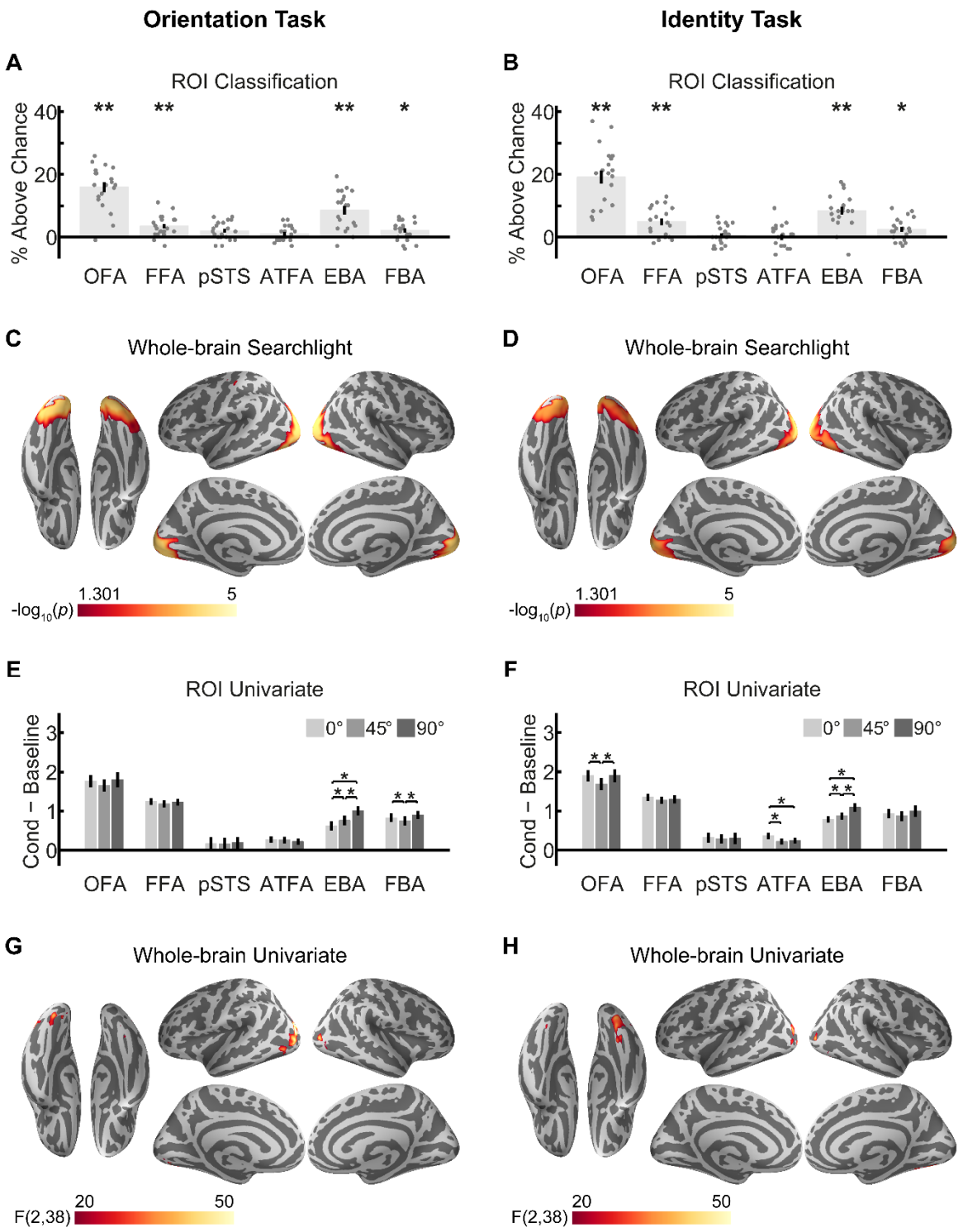

Figure 3. Neural responses to face orientation. (A) and (B) show classification of face orientation above chance-level in ROIs. Grey scatter points show classification accuracies for individual participants, ** indicates $p < 0.001$, * indicates $p < 0.05$, Bonferroni corrected for N = 6 ROIs. (C) and (D) show classification of face orientation in whole-brain searchlight analyses. The scale bar shows $-\log_{10}(p$ values) between 1.301 ($p = 0.05$) and 5 ($p = 1 \times 10^{-5}$), FWE corrected. (E) and (F) show differences in mean BOLD activation to faces of different orientations in ROIs. * indicates $p < 0.05$.

(G) and (H) show differences in mean BOLD activation to faces of different orientations in a whole-brain analysis (FWE corrected). All analyses were conducted separately for fMRI data collected while participants responded to stimulus orientation (A), (C), (E) and (G) or to stimulus identity (B), (D), (F) and (H).

Classification of face orientation from the orientation task data (Fig. 3A) was significantly above chance-level (33 ⅓ %) in the face-responsive OFA (49.3 %, $p < 0.0006$ Bonferroni corrected, Cohen's $d = 2.31$) and FFA (36.9 %, $p < 0.0006$ Bonferroni corrected, Cohen's $d = 1.05$), but not in the pSTS (35.4 %, $p = 0.014$ uncorrected, Cohen's $d = 0.70$) or the ATFA (34.4 %, $p = 0.12$ uncorrected, Cohen's $d = 0.43$). Classification of face orientation from the orientation task data was also significantly above chance in both body-responsive ROIs (EBA: 41.9%, $p < 0.0006$ Bonferroni corrected, Cohen's $d = 1.42$; FBA: 35.5%, $p = 0.046$ Bonferroni corrected, Cohen's $d = 0.67$). These results were identical for classification of face orientation from the identity task data (Fig. 3B). Classification was significantly above chance in the face-responsive OFA (52.5 %, $p < 0.0006$ Bonferroni corrected, Cohen's $d = 2.05$) and FFA (38.2 %, $p < 0.0006$ Bonferroni corrected, Cohen's $d = 1.06$) and the body-responsive EBA (41.8 %, $p < 0.0006$ Bonferroni corrected, Cohen's $d = 1.56$) and FBA (35.9 %, $p = 0.020$ Bonferroni corrected, Cohen's $d = 0.71$), but was not higher than chance in the pSTS (33.9 %, $p = 0.25$ uncorrected, Cohen's $d = 0.19$) or ATFA (33.3 %, $p = 0.52$ uncorrected, Cohen's $d = -0.01$).

We performed whole-brain searchlight analyses to investigate if any other regions could classify face orientation (Fig 3C-D). Using fMRI data from both the orientation and identity tasks, we found that face orientation could be decoded from a large area of occipitotemporal cortex, including face- and body-responsive regions as well as the early visual cortex. Consistent with our ROI results, this region included the face-responsive OFA and FFA and the body-responsive EBA and FBA. From the orientation task fMRI data we could also decode orientation in the left motor cortex as participants' responded using button presses with different fingers to indicate which orientation was shown in the block.

4.3.2.2. Univariate analyses

We investigated if there were differences in the overall level of neural activity to faces of different orientations. To do this, we conducted one-way repeated measures ANOVAs with 3 levels (0°, 45° and 90°) in our face- and body-responsive ROIs as well as in whole-brain analyses. We conducted the analyses separately for fMRI data from the two behavioural tasks (orientation and identity tasks). The results are shown in Fig. 3E-H.

For fMRI data from the orientation task (Fig. 3E), we found a significant effect of face orientation condition in the body-responsive EBA ($F_{2,38}$ = 67.76, p = 1.76 x 10^{-12} Bonferroni corrected, η_p^2 = 0.78) and the FBA ($F_{2,38}$ = 10.01, p = 0.0019 Bonferroni corrected, η_p^2 = 0.35), but no effect in any of the face-responsive ROIs (OFA: $F_{2,38}$ = 3.45, p = 0.042 uncorrected, η_p^2 = 0.15; FFA: $F_{2,38}$ = 1.39, p = 0.26 uncorrected, η_p^2 = 0.068; pSTS: $F_{2,38}$ = 0.27, p = 0.76 uncorrected, η_p^2 = 0.014; ATFA: $F_{2,38}$ = 0.81, p = 0.45 uncorrected, η_p^2 = 0.041). Follow-up paired *t*-tests showed that in the EBA there was higher BOLD activation to 90° faces compared to 45° faces (M = 0.24, SE = 0.029, t_{19} = 8.41, p = 7.97 x 10^{-8}, Cohen's d = 1.88) and 0° faces (M = 0.38, SE = 0.035, t_{19} = 11.01, p = 1.09 x 10^{-9}, Cohen's d = 2.46), and higher BOLD activation to 45° faces compared to 0° faces (M = 0.14, SE = 0.036, t_{19} = 3.86, p = 0.0011, Cohen's d = 0.86). In the FBA, follow-up paired *t*-tests showed that BOLD activation was lower for 45° faces compared to 0° faces (M = -0.078, SE = 0.035, t_{19} = -2.26, p = 0.036, Cohen's d = -0.50) and 90° faces (M = -0.15, SE = 0.028, t_{19} = -5.22, p = 4.90 x 10^{-5}, Cohen's d = -1.17), but there was no difference in BOLD activation between 0° and 90° faces (M = -0.070, SE = 0.036, t_{19} = -1.95, p = 0.066, Cohen's d = -0.44).

For fMRI data from the identity task (Fig. 3F), we found a significant effect of face orientation condition in the face-responsive OFA ($F_{2,38}$ = 9.01, p = 0.0038 Bonferroni corrected, η_p^2 = 0.32) and ATFA ($F_{2,38}$ = 7.77, p = 0.023 Bonferroni corrected and Greenhouse-Geisser corrected for non-sphericity, η_p^2 = 0.29), but not in the FFA ($F_{2,38}$ = 2.02, p = 0.15 uncorrected, η_p^2 = 0.096) or pSTS ($F_{2,38}$ = 0.28, p = 0.76 uncorrected, η_p^2 = 0.015). There was also a significant effect of face orientation condition in the body-responsive EBA ($F_{2,38}$ = 41.01, p = 1.79 x 10^{-6} Bonferroni corrected and Greenhouse-Geisser corrected for non-sphericity, η_p^2 = 0.68), but not in the FBA ($F_{2,38}$ = 5.20, p = 0.010 uncorrected, η_p^2 = 0.21). Follow-up paired *t*-tests in the OFA showed that there was lower BOLD activation to

45° faces compared to 0° faces (M = -0.21, SE = 0.055, t_{19} = -3.85, p = 0.0011, Cohen's d = -0.86) and 90° faces (M = -0.22, SE = 0.057, t_{19} = -3.77, p = 0.0013, Cohen's d = -0.84) but no difference between 0° and 90° faces (M = -0.0056, SE = 0.062, t_{19} = -0.091, p = 0.93, Cohen's d = -0.020). In the ATFA, BOLD activation was higher for 0° faces compared to 45° (M = 0.15, SE = 0.047, t_{19} = 3.16, p = 0.0052, Cohen's d = 0.71) and 90° (M = 0.13, SE = 0.045, t_{19} = 2.84, p = 0.011, Cohen's d = 0.63) faces, and no difference in activation between 45° and 90° faces (M = -0.020, SE = 0.027, t_{19} = -0.74, p = 0.47, Cohen's d = -0.17). Similarly to the fMRI data from the orientation task, in the EBA there was higher BOLD activation to 90° faces compared to 45° faces (M = 0.22, SE = 0.027, t_{19} = 8.01, p = 1.66 x10^{-7}, Cohen's d = 1.79) and 0° faces (M = 0.31, SE = 0.046, t_{19} = 6.64, p = 2.39 x10^{-6}, Cohen's d = 1.48), and higher BOLD activation to 45° faces compared to 0° faces (M = 0.089, SE = 0.028, t_{19} = 3.19, p = 0.0048, Cohen's d = 0.71).

We performed a whole-brain analysis to see if any other brain regions would show differences in mean BOLD activation to faces of different orientations (Fig 3G-H). We identified several clusters in the occipital and fusiform cortices, which overlapped with the locations of the EBA, OFA and early visual cortex and slightly overlapped with the FFA and FBA.

4.3.3. Neural responses to body orientation

4.3.3.1. Classification analyses

We investigated which brain regions have different patterns of neural activity evoked by different body orientations that could generalize across body identity. We trained a linear SVM to distinguish between patterns of neural activity evoked by the three body orientations, using two body identities for training the classifier. We then tested the trained classifier on its ability to classify body orientation from a third identity in a left out run of fMRI data. Again, we used a leave one run out cross-validation procedure and also used each body identity once as the left out test identity. As previously, we performed the analyses in face- and body-responsive ROIs and whole-brain searchlights, and repeated the analysis for fMRI data from the two behavioural tasks (orientation and identity tasks). The results are shown in Fig. 4A-D.

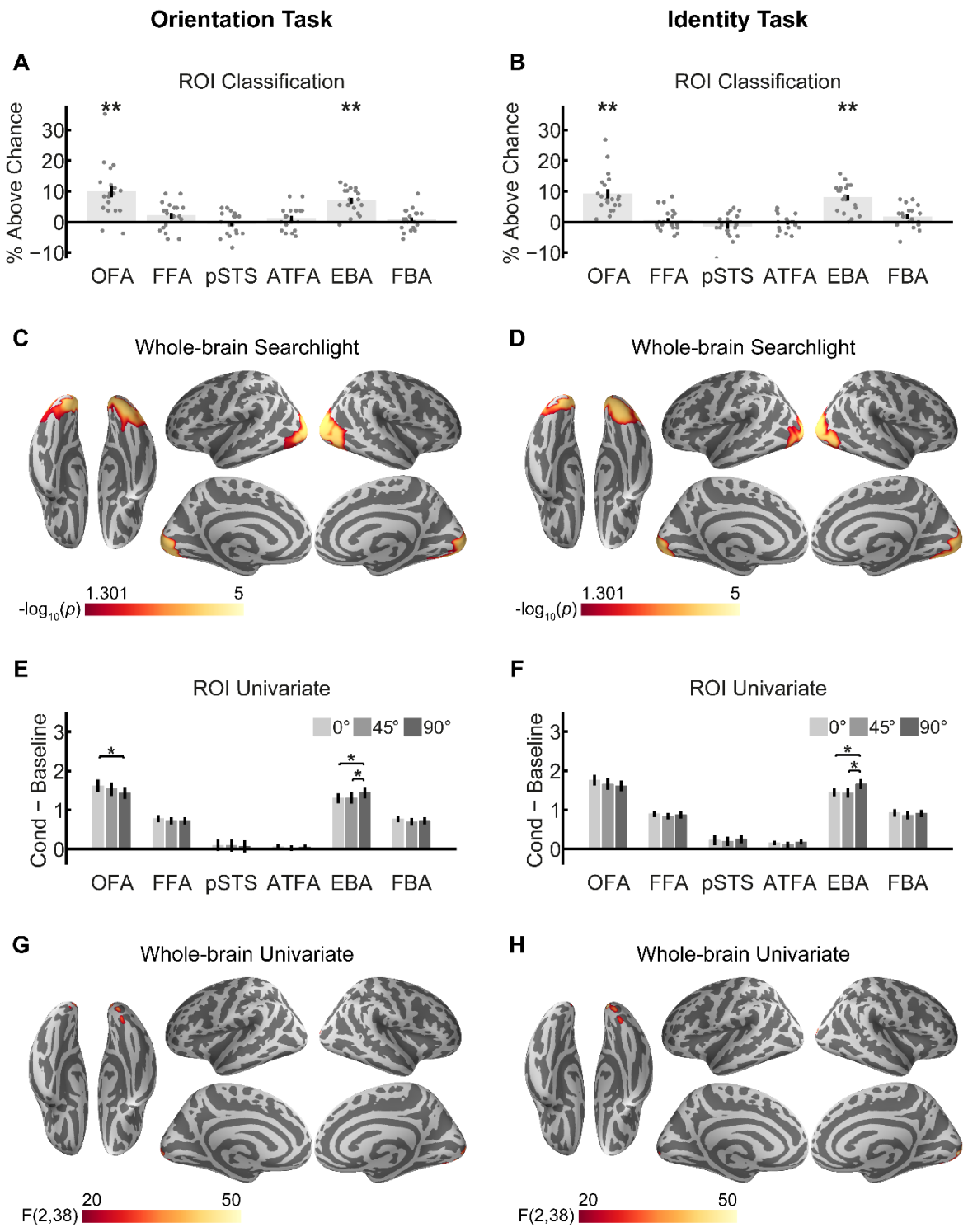

Figure 4. Neural responses to body orientation. (A) and (B) show classification of body orientation above chance-level in ROIs. Grey scatter points show classification accuracies for individual participants, ** indicates $p < 0.001$, Bonferroni corrected for N = 6 ROIs. (C) and (D) show classification of body orientation in whole-brain searchlight analyses. The scale bar shows $-\log_{10}(p$ values) between 1.301 ($p = 0.05$) and 5 ($p = 1 \times 10^{-5}$), FWE corrected. (E) and (F) show differences in mean BOLD activation to bodies of different orientations in ROIs. * indicates $p < 0.05$. (G) and (H)

show differences in mean BOLD activation to bodies of different orientations in a whole-brain analysis (FWE corrected). All analyses were conducted separately for fMRI data collected while participants responded to stimulus orientation (A), (C), (E) and (G) or to stimulus identity (B), (D), (F) and (H).

For the orientation task data (Fig. 4A), classification of body orientation was significantly above chance-level (33 ⅓ %) in the body-responsive EBA (40.3 %, $p < 0.0006$ Bonferroni corrected, Cohen's $d = 1.79$) and the face-responsive OFA (43.3 %, $p < 0.0006$ Bonferroni corrected, Cohen's $d = 1.17$). Classification of body orientation was not significantly higher than chance in the body-responsive FBA (34.0 %, $p = 0.21$ uncorrected, Cohen's $d = 0.20$) or face-responsive FFA (35.4 %, $p = 0.010$ uncorrected, Cohen's $d = 0.47$), pSTS (32.8 %, $p = 0.72$ uncorrected, Cohen's $d = -0.14$) or ATFA (34.5 %, $p = 0.086$ uncorrected, Cohen's $d = 0.32$). Results were identical for the identity task data (Fig. 4B). Classification of body orientation was significantly above chance in the EBA (41.3 %, $p < 0.0006$ Bonferroni corrected, Cohen's $d = 1.77$) and OFA (42.5 %, $p < 0.0006$ Bonferroni corrected, Cohen's $d = 1.43$), but not in any other ROIs we tested (FBA: 35.0 %, $p = 0.029$ uncorrected, Cohen's $d = 0.48$; FFA: 33.8 %, $p = 0.29$ uncorrected, Cohen's $d = 0.14$; pSTS: 32.0 %, $p = 0.93$ uncorrected, Cohen's $d = -0.33$; ATFA: 33.1 %, $p = 0.63$ uncorrected, Cohen's $d = -0.10$).

We performed whole-brain searchlight analyses to investigate if any other regions could classify body orientation (Fig 4C-D). We found a large area of occipitotemporal cortex, including the early visual cortex, could decode body orientation. For fMRI data from the orientation task this region overlapped with the OFA and EBA, and slightly overlapped with the FFA and rFBA. For fMRI data from the identity task this region overlapped with the OFA and EBA, and slightly overlapped with the FFA. In fMRI data from the orientation task we could also decode orientation in the left motor cortex as participants' responded using button presses with different fingers to indicate which orientation was shown in the block.

4.3.3.2. Univariate analyses

We investigated if there were differences in the overall level of neural activity to bodies of different orientations. We conducted one-way repeated measures ANOVAs with 3

levels (0°, 45° and 90°) in both ROI and whole-brain analyses, separately for fMRI data from the two behavioural tasks (orientation and identity tasks). The results are shown in Fig. 4E-H.

For fMRI data from the orientation task (Fig. 4E), we found significant differences in BOLD response to body orientation in the body-responsive EBA ($F_{2,38}$ = 6.06, p = 0.031 Bonferroni corrected, η_p^2 = 0.24) and face-responsive OFA ($F_{2,38}$ = 6.09, p = 0.031 Bonferroni corrected, η_p^2 = 0.24), but not in any other ROI (FBA: $F_{2,38}$ = 2.06, p = 0.14 uncorrected, η_p^2 = 0.098; FFA: $F_{2,38}$ = 1.56, p = 0.22 uncorrected, η_p^2 = 0.076; pSTS: $F_{2,38}$ = 0.35, p = 0.70, η_p^2 = 0.018; ATFA: $F_{2,38}$ = 0.30, p = 0.74 uncorrected, η_p^2 = 0.016). Follow-up paired t-tests in the EBA showed that this result was driven by higher activity to 90° bodies compared to 0° (M = 0.14, SE = 0.050, t_{19} = 2.82, p = 0.011, Cohen's d = 0.63) and 45° (M = 0.13, SE = 0.044, t_{19} = 3.05, p = 0.0065, Cohen's d = 0.68) bodies, but there was no difference in activation between 0° and 45° bodies (M = -0.0078, SE = 0.043, t_{19} = -0.18, p = 0.86, Cohen's d = -0.041). In the OFA, follow up paired t-tests showed there was higher activity to 0° compared to 90° bodies (M = 0.18, SE = 0.051, t_{19} = 3.51, p = 0.0023, Cohen's d = 0.79), but no differences in activation between 0° and 45° bodies (M = 0.085, SE = 0.047, t_{19} = 1.81, p = 0.086, Cohen's d = 0.41) or between 45° and 90° bodies (M = 0.095, SE = 0.056, t_{19} = 1.69, p = 0.11, Cohen's d = 0.38).

For fMRI data from the identity task (Fig. 4F), we found significant differences in BOLD response to bodies of different orientations in the EBA ($F_{2,38}$ = 12.5, p = 4.12 x 10^{-4} Bonferroni corrected, η_p^2 = 0.40), but not in any other ROIs we tested (FBA: $F_{2,38}$ = 0.95, p = 0.40 uncorrected, η_p^2 = 0047; OFA: $F_{2,38}$ = 4.41, p = 0.019 uncorrected, η_p^2 = 0.19; FFA: $F_{2,38}$ = 1.35, p = 0.27 uncorrected, η_p^2 = 0.067; pSTS: $F_{2,38}$ = 0.77, p = 0.47 uncorrected, η_p^2 = 0.039; ATFA: $F_{2,38}$ = 1.08, p = 0.35 uncorrected, η_p^2 = 0.054). Follow-up paired t-tests in the EBA showed that, as for the orientation task fMRI data, there was higher activity to 90° bodies compared to 0° (M = 0.21, SE = 0.048, t_{19} = 4.32, p = 3.70 x 10^{-4}, Cohen's d = 0.97) and 45° (M = 0.22, SE = 0.049, t_{19} = 4.51, p = 2.41 x 10^{-4}, Cohen's d = 1.01) bodies, but no difference in activation between 0° and 45° bodies (M = 0.016, SE = 0.052, t_{19} = 0.30, p = 0.76, Cohen's d = 0.068).

We performed a whole-brain analysis to see if any other brain regions would show differences in mean BOLD activation to bodies of different orientations (Fig 4G-H). From these analyses we found bilateral clusters in the EVC from fMRI data from both the orientation and identity behavioural tasks. In the data from the identity task, we identified additional clusters that overlapped with the EBA.

4.3.4. Classification of orientation across face and body stimuli

We investigated whether the patterns of neural activity evoked by stimuli of different orientations could abstract across neural activity evoked by face and body stimuli. To do this, we trained a linear SVM classifier to distinguish between patterns of neural activity evoked by faces of different orientations and then tested this trained classifier on its ability to classify the orientation of neural activity evoked by body stimuli (and vice-versa using neural activity evoked by body stimuli for training the classifier and neural activity evoked by face stimuli for testing it). As previously, we used a leave one run out cross-validation method to ensure separation of training and testing data. We performed the analysis in face- and body-responsive ROIs and whole-brain searchlight analyses, separately for fMRI data from the two behavioural tasks (orientation and identity tasks). The results are shown in Figure 5.

For the orientation task data (Fig. 5A), classification of orientation across neural activity evoked by face and body stimuli was significantly above chance-level (33 ⅓ %) in the OFA (38.6 %, $p < 0.0006$ Bonferroni corrected, Cohen's $d = 1.23$) and the EBA (37.2 %, $p < 0.0006$ Bonferroni corrected, Cohen's $d = 1.18$), but not in any other face-responsive ROIs (FFA: 34.8 %, $p = 0.046$ uncorrected, Cohen's $d = 0.61$; pSTS: 34.8 %, $p = 0.053$ uncorrected, Cohen's $d = 0.61$; ATFA: 33.8 %, $p = 0.32$ uncorrected, Cohen's $d = 0.17$) or in the FBA (33.9 %, $p = 0.26$ uncorrected, Cohen's $d = 0.26$). These results were identical for analyses using fMRI data from the identity task (Fig. 5B). Classification was significantly above chance in the OFA (37.9 %, $p < 0.0006$ Bonferroni corrected, Cohen's $d = 1.00$) and the EBA (37.7 %, $p < 0.0006$ Bonferroni corrected, Cohen's $d = 1.30$), but not in the other ROIs tested (FFA: 34.9 %, $p = 0.038$ uncorrected, Cohen's $d = 0.48$; pSTS: 34.5 %, $p = 0.081$ uncorrected, Cohen's $d = 0.38$; ATFA: 33.6 %, $p = 0.39$ uncorrected, Cohen's $d = 0.095$; FBA: 34.9 %, $p = 0.042$ uncorrected, Cohen's $d = 0.50$).

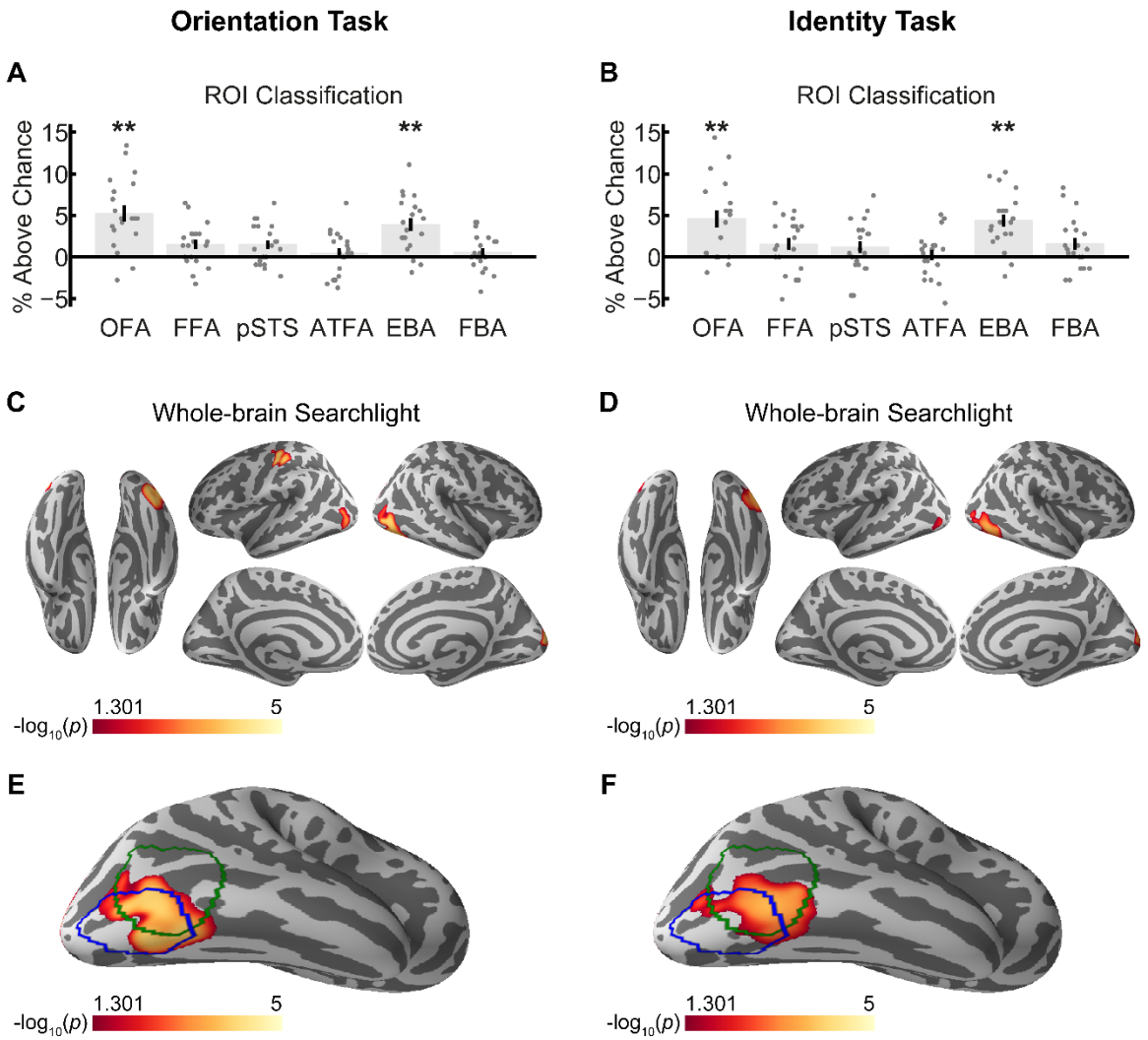

Figure 5. Classification of orientation across neural activity evoked by face and body stimuli. (A) and (B) show classification of orientation above chance-level in ROIs. Grey scatter points show classification accuracies for individual participants, ** indicates $p < 0.001$, Bonferroni corrected for N = 6 ROIs. (C), (D), (E) and (F) show classification of orientation in whole-brain searchlight analyses. Scale bars show $-\log_{10}(p$ values) between 1.301 ($p = 0.05$) and 5 ($p = 1 \times 10^{-5}$), FWE corrected. (E) and (F) show the position of the right hemisphere searchlight clusters compared to the mean locations of the EBA (in green) and OFA (in blue). All analyses were conducted separately for fMRI data collected while participants responded to stimulus orientation (A), (C) and (E) or to stimulus identity (B), (D) and (F).

We performed whole-brain searchlight analyses to investigate if any other regions could classify orientation across neural activity evoked by face and body stimuli (Fig 5C-F). Consistent with our ROI analyses, we found bilateral regions overlapping with the OFA and EBA that were able to decode orientation. Interestingly these regions were located at the intersection of the OFA and EBA (see Fig 5E-F). Orientation could also be decoded in the right early visual cortex. In fMRI data from the orientation task we could also decode orientation in the left motor cortex, due to participants' button presses in this task.

As we found some differences in the mean BOLD responses to faces and bodies of different orientations (Fig 3E-F and Fig 4E-F), we performed a control analysis to ensure our classification results across neural activity evoked by face and body stimuli were not driven by differences in the mean BOLD signal. We repeated these classification analyses using only the mean BOLD signal in each ROI, or in each searchlight sphere for training and testing the classifier. If differences in mean BOLD activation were driving the classification results, there should be identical classification results in these analyses. Results showed that we could not decode orientation in any ROI using the mean BOLD activation (Fig S1A-B). Furthermore, searchlight results showed a cluster in the right EVC, but not in any other brain region (Fig S1C-D). These results suggest that our decoding of orientation in the right EVC may be driven by differences in the mean BOLD signal, but we find no evidence that activity in EBA or OFA was affected by any differences in the overall mean signal.

4.4. Discussion

In this study, we investigated the neural responses to faces and bodies varying in orientation. We show that the OFA, EBA and early visual cortex contain different patterns of activity to both faces and bodies of different orientations. Furthermore, orientation responses in the OFA and EBA were abstract. In these regions, a classifier trained to distinguish patterns of neural activity evoked by face orientations could then decode orientation form neural activity evoked by bodies, and vice-versa. We also show that the FFA and FBA respond to face orientation but not to body orientation, suggesting that orientation responses in these areas are face-specific. These results show that there are both similarities and differences in the neural processing of face and body orientation.

4.4.1. Orientation responses in the OFA and EBA

We found that we could consistently decode both face and body orientation from the face-responsive OFA and body-responsive EBA. Several studies have shown that the OFA responds to face orientation (Axelrod & Yovel, 2012; Flack, Harris, Young, & Andrews, 2019; Guntupalli et al., 2016; Kietzmann et al., 2012), but no previous study, to our knowledge, has shown that it responds to body orientation. Similarly, previous studies have shown that the EBA responds to body orientation (Chan et al., 2004; Ewbank et al., 2011; Taylor et al., 2010), but no previous study has shown that it responds to face orientation. We also found some differences in the univariate activation to both faces and bodies of different orientations in the OFA and EBA. In the OFA, we found lower responses to 45° faces compared to 0° and 90° faces and higher responses to 0° bodies compared to 90°bodies. In the EBA, we found responses progressively increased from 0° to 90° faces and higher responses to 90° bodies compared to 0° and 45° bodies, suggesting a preference for profile views in the EBA. In sum, these results show that orientation responses in the OFA and EBA are not face or body selective.

We further investigated if patterns of neural responses to orientation in the OFA and EBA could generalize across neural activity evoked by faces and bodies. In both the OFA and EBA we found that classifiers trained to distinguish patterns of neural activity evoked by different face orientations could later decode patterns of neural activity evoked by body

orientations, and vice-versa. Furthermore, our searchlight analyses showed that there was strong classification at the intersection of the OFA and EBA (Fig. 5E-F). These results suggest that the OFA and EBA contain an abstract high-level encoding of person orientation, as faces and bodies vary considerably in their low-level visual features. Interestingly, two recent studies have shown that the EBA also responds higher to interacting people as compared to non-interacting people (Abassi & Papeo, 2019; Walbrin & Koldewyn, 2019). In addition, a region in nearby occipital cortex has been identified that is involved in gaze following behaviour (Marquardt, Ramezanpour, Dicke, & Thier, 2017). These findings, in combination with our results, suggest that person orientation processing may be an important function of this region of occipital cortex.

4.4.2. Face orientation responses in the FFA, FBA and ATFA

Consistent with previous studies, we were able to decode face orientation in the FFA (Axelrod & Yovel, 2012; Guntupalli et al., 2016; Kietzmann et al., 2012; Ramírez et al., 2014). We could also decode face orientation from the body-responsive FBA, however, as this region is known to overlap with the FFA (Schwarzlose, Baker, & Kanwisher, 2005), it is possible that the overlapping voxels from FFA may have contributed to our face orientation classification in the FBA. In contrast to our face orientation results, we were unable to decode body orientation from the either the FFA or FBA. This result shows a difference to previous work showing a sensitivity to body orientation in the FBA (Ewbank et al., 2011; Taylor et al., 2010). This difference may be due to experimental design as the previous studies used repetition suppression designs, whereas in this study we used a MVPA approach. Interestingly, more body orientation information in the EBA than the FBA mirrors a recent finding in macaque monkeys (Kumar et al., 2017). The authors found that middle STS body patch contained more body orientation information than the anterior STS body patch. Finally, we note that face but not body orientation responses in the fusiform gyrus is consistent with a previous finding that orientation responses in FFA are face-specific (Ramírez et al., 2014).

Although we were unable to decode orientation from the ATFA, we found higher neural responses to frontal faces compared to 45° or 90° ones in this area. This finding shows similarity to work in macaques demonstrating that there are a higher proportion of

neurons responding to frontal faces as compared to other orientations in the most anterior face-responsive patch (AM) (Dubois et al., 2015; Freiwald & Tsao, 2010). Thus, our result in ATFA supports a homology between the human ATFA and macaque AM.

4.4.3. Superior temporal sulcus (STS)

The superior temporal sulcus is thought to be involved in processing changeable aspects of faces, such as emotional expression and gaze (Haxby, Hoffman, & Gobbini, 2000). fMRI studies have found mixed results as to the involvement of this area in representing face orientation (Axelrod & Yovel, 2012; Ramírez, 2018; Ramírez et al., 2014). However, neurons have been identified in macaque anterior STS that respond to orientation in an abstract manner across face and body stimuli (Wachsmuth et al., 1994). In humans, the anterior STS has also been shown to respond to gaze direction invariant to head orientation (Carlin, Calder, Kriegeskorte, Nili, & Rowe, 2011). We did not find evidence for face or body orientation coding in either our pSTS ROI, or any region of the STS in our searchlight analyses. However, we note that classification of face orientation and classification of orientation across face and body stimuli was close to significance in the pSTS when participants were performing the orientation task. This suggests there may be orientation information present in the pSTS, which failed to reach the threshold for significance in this study. We note that the selective involvement of the STS while participants attended to orientation would also be consistent with this region's known role in gaze processing (Carlin & Calder, 2013).

4.4.4. Similarities and differences across behavioural task

Participants performed two different behavioural tasks in this study while we recorded their brain activity using fMRI. In one half of the dataset participants responded to stimulus orientation, whereas in the other half of the dataset they responded to stimulus identity. Thus, we were able to investigate if there would be differences in neural coding based on whether participants attended to stimulus orientation or identity. Interestingly, we found very few differences in our results from this task modulation. In fact, our results demonstrate considerable consistency as they largely replicate across the two halves of our dataset. One interesting difference is that we found differences in face orientation responses in the ATFA only when participants performed the identity task. The ATFA is

known to be involved in processing identity (Anzellotti, Fairhall, & Caramazza, 2014; Guntupalli et al., 2016; Kriegeskorte, Formisano, Sorger, & Goebel, 2007; Nasr & Tootell, 2012; Nestor, Plaut, & Behrmann, 2011), thus we hypothesize that face orientation response differences in this region may only be identifiable when it is engaged in a task that optimally drives its neural responses.

4.4.5. Conclusion

We show that a region in the occipital cortex, located at the intersection of the OFA and EBA, contains patterns of neural activity evoked by orientation that can generalize across neural activity evoked by faces and bodies. As faces and bodies vary considerably in their low-level properties, this result suggests that this region responds to person orientation in a high-level abstract code. Furthermore, we show that regions in the fusiform gyrus (FFA and FBA) respond to face but not body orientation, suggesting that responses to face orientation are more distributed than those to body orientation. Our results offer new insights into how, and where in the brain, person orientation information is encoded.

Acknowledgements

This research was supported by the Max Planck Society, Germany.

References

Abassi, E., & Papeo, L. (2019). Seeing relationships: The specialization for a two-body shape in human visual perception. *BioRxiv*, (Advance online publication, retrieved 6th August 2019), 637082. https://doi.org/10.1101/637082

Anzellotti, S., Fairhall, S. L., & Caramazza, A. (2014). Decoding Representations of Face Identity That are Tolerant to Rotation. *Cerebral Cortex*, *24*(8), 1988–1995. https://doi.org/10.1093/cercor/bht046

Axelrod, V., & Yovel, G. (2012). Hierarchical Processing of Face Viewpoint in Human Visual Cortex. *The Journal of Neuroscience*, *32*(7), 2442–2452. https://doi.org/10.1523/JNEUROSCI.4770-11.2012

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Brooks, J. L. (2012). Counterbalancing for serial order carryover effects in experimental condition orders. *Psychological Methods*, *17*(4), 600–614. https://doi.org/10.1037/a0029310

Carlin, J. D., & Calder, A. J. (2013). The neural basis of eye gaze processing. *Current Opinion in Neurobiology*, *23*(3), 450–455. https://doi.org/10.1016/j.conb.2012.11.014

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. https://doi.org/10.1016/j.cub.2011.09.025

Chan, A. W.-Y., Peelen, M. V., & Downing, P. E. (2004). The effect of viewpoint on body representation in the extrastriate body area. *NeuroReport*, *15*(15), 2407–2410. https://doi.org/10.1097/00001756-200410250-00021

Dubois, J., de Berker, A. O., & Tsao, D. Y. (2015). Single-Unit Recordings in the Macaque Face Patch System Reveal Limitations of fMRI MVPA. *The Journal of Neuroscience*, *35*(6), 2791–2802. https://doi.org/10.1523/JNEUROSCI.4037-14.2015

Ewbank, M. P., Lawson, R. P., Henson, R. N., Rowe, J. B., Passamonti, L., & Calder, A. J. (2011). Changes in "Top-Down" Connectivity Underlie Repetition Suppression in the Ventral Visual Pathway. *The Journal of Neuroscience*, *31*(15), 5635–5642. https://doi.org/10.1523/JNEUROSCI.5013-10.2011

Flack, T. R., Harris, R. J., Young, A. W., & Andrews, T. J. (2019). Symmetrical Viewpoint Representations in Face-Selective Regions Convey an Advantage in the Perception and Recognition of Faces. *The Journal of Neuroscience*, *39*(19), 3741–3751. https://doi.org/10.1523/jneurosci.1977-18.2019

Freiwald, W. A., & Tsao, D. Y. (2010). Functional Compartmentalization and Viewpoint Generalization Within the Macaque Face-Processing System. *Science*, *330*(6005), 845–851.

Gibson, J. J., & Pick, A. D. (1963). Perception of Another Person's Looking Behavior. *The American Journal of Psychology*, *76*(3), 386–394.

Guntupalli, J. S., Wheeler, K. G., & Gobbini, M. I. (2016). Disentangling the Representation of Identity from Head View Along the Human Face Processing Pathway. *Cerebral Cortex*, *27*(1), 46–53. https://doi.org/10.1093/cercor/bhw344

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*(6), 223–233. https://doi.org/10.1016/S1364-6613(00)01482-0

Hebart, M. N., Görgen, K., & Haynes, J.-D. (2015). The Decoding Toolbox (TDT): a versatile software package for multivariate analyses of functional imaging data. *Frontiers in Neuroinformatics*, *8*, 1–18. https://doi.org/10.3389/fninf.2014.00088

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

11772. https://doi.org/10.1523/jneurosci.0126-12.2012

Kleiner, M., Brainard, D., & Pelli, D. (2007). "What's new in Psychtoolbox-3?" In *Perception 36 ECVP Abstract Supplement.*

Kriegeskorte, N., Formisano, E., Sorger, B., & Goebel, R. (2007). Individual faces elicit distinct response patterns in human anterior temporal cortex. *Proceedings of the National Academy of Sciences*, *104*(51), 20600–20605. https://doi.org/10.1073/pnas.0705654104

Kumar, S., Popivanov, I. D., & Vogels, R. (2017). Transformation of Visual Representations Across Ventral Stream Body-selective Patches. *Cerebral Cortex*, *29*(1), 215–229. https://doi.org/10.1093/cercor/bhx320

Li, T., Bolkart, T., Black, M. J., Li, H., & Romero, J. (2017). Learning a model of facial shape and expression from 4D scans. *ACM Transactions on Graphics*, *36*(6), 1–17. https://doi.org/10.1145/3130800.3130813

Loper, M., Mahmood, N., Romero, J., Pons-moll, G., & Black, M. J. (2015). SMPL : A Skinned Multi-Person Linear Model. *ACM Trans. Graphics (Proc. SIGGRAPH Asia)*, *34*(6), 248:1--248:16. https://doi.org/10.1145/2816795.2818013

Marquardt, K., Ramezanpour, H., Dicke, P. W., & Thier, P. (2017). Following Eye Gaze Activates a Patch in the Posterior Temporal Cortex That Is not Part of the Human "Face Patch" System. *Eneuro*, *4*(2), 1–10. https://doi.org/10.1523/eneuro.0317-16.2017

Moors, P., Germeys, F., Pomianowska, I., & Verfaillie, K. (2015). Perceiving where another person is looking: The integration of head and body information in estimating another person's gaze. *Frontiers in Psychology*, *6*(909), 1–12. https://doi.org/10.3389/fpsyg.2015.00909

Nasr, S., & Tootell, R. B. H. (2012). Role of fusiform and anterior temporal cortical areas in facial recognition. *NeuroImage*, *63*(3), 1743–1753. https://doi.org/10.1016/j.neuroimage.2012.08.031

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. https://doi.org/10.1162/jocn.2009.21312.

Nestor, A., Plaut, D. C., & Behrmann, M. (2011). Unraveling the distributed neural code of facial identity through spatiotemporal pattern analysis. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(24), 9998–10003. https://doi.org/10.1073/pnas.1102433108

Nichols, T. E., & Holmes, A. P. (2001). Nonparametric Permutation Tests For Functional Neuroimaging: A Primer with Examples. *Human Brain Mapping*, *15*(1), 1–25. https://doi.org/10.1002/hbm.1058

Perrett, D. I., Smith, P. A. J., 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, Biological Sciences*, *223*(1232), 293–317.

Ramírez, F. M. (2018). Orientation Encoding and Viewpoint Invariance in Face Recognition: Inferring Neural Properties from Large-Scale Signals. *Neuroscientist*, *24*(6), 582–608. https://doi.org/10.1177/1073858418769554

Ramírez, F. M., Cichy, R. M., Allefeld, C., & Haynes, J.-D. (2014). The neural code for face orientation in the human fusiform face area. *The Journal of Neuroscience*, *34*(36), 12155–12167. https://doi.org/10.1523/JNEUROSCI.3156-13.2014

Schwarzlose, R. F., Baker, C. I., & Kanwisher, N. (2005). Separate Face and Body Selectivity on the Fusiform Gyrus. *The Journal of Neuroscience*, *25*(47), 11055–11059. https://doi.org/10.1523/JNEUROSCI.2621-05.2005

Taylor, J. C., Wiggett, A. J., & Downing, P. E. (2010). fMRI-adaptation studies of viewpoint tuning in the extrastriate and fusiform body areas. *Journal of Neurophysiology*, *103*(3), 1467–1477. https://doi.org/10.1152/jn.00637.2009

Wachsmuth, E., Oram, M. W., & Perrett, D. I. (1994). Recognition of objects and their component parts:

Responses of single units in the temporal cortex of the macaque. *Cerebral Cortex*, *4*(5), 509–522. https://doi.org/10.1093/cercor/4.5.509

Walbrin, J., & Koldewyn, K. (2019). Dyadic interaction processing in the posterior temporal cortex. *NeuroImage*, *198*, 296–302. https://doi.org/10.1016/j.neuroimage.2019.05.027

Wollaston, W. H. (1824). On the Apparent Direction of Eyes in a Portrait. *Philosophical Transactions of the Royal Society of London*, *114*, 247–256.

Supplemental Information

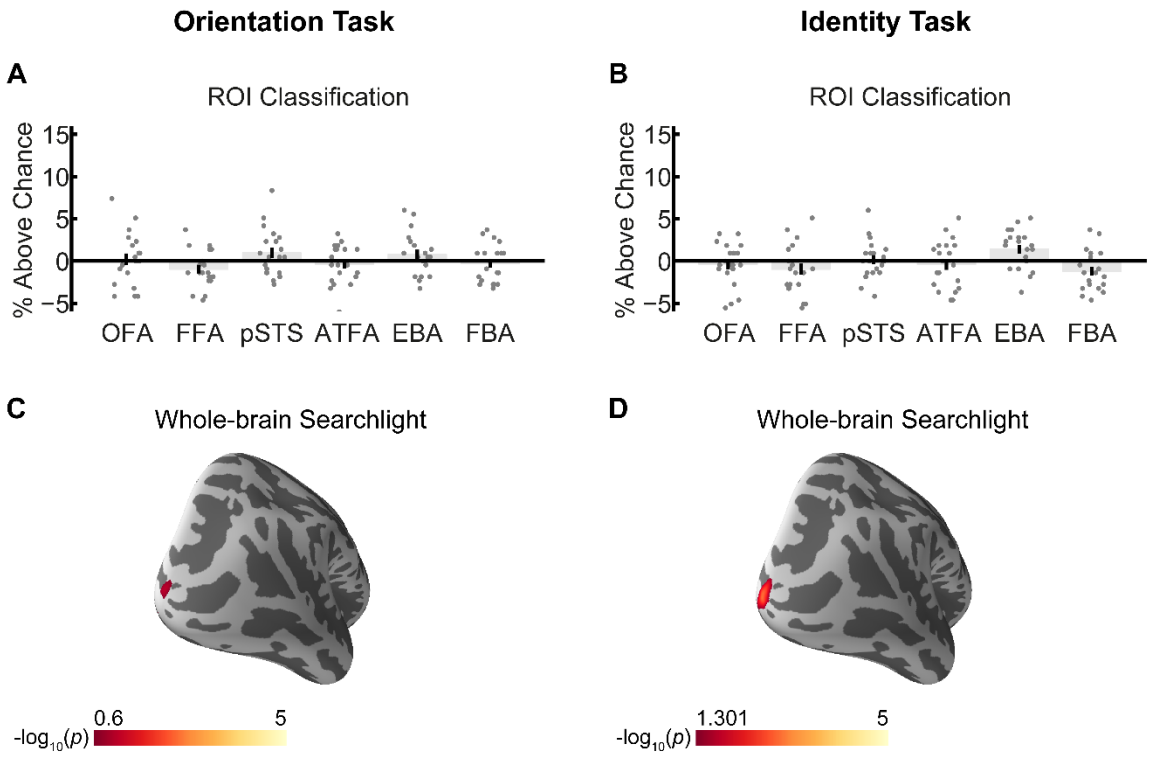

Figure S1. Classification of orientation across neural activity evoked by face and body stimuli using the mean BOLD signal. (A) and (B) show classification of orientation in ROIs. Grey scatter points show classification accuracies for individual participants. (C) and (D) show classification of orientation in whole-brain searchlight analyses. For (C) the scale bar shows $-\log_{10}(p$ values) between 0.6 (p = 0.25) and 5 (p = 1 x10^{-5}), FWE corrected. For (D) the scale bar shows $-\log_{10}(p$ values) between 1.301 (p = 0.05) and 5 (p = 1 x10^{-5}), FWE corrected. All analyses were conducted separately for fMRI data collected while participants responded to stimulus orientation (A) and (C) or to stimulus identity (B) and (D).

5. Investigating holistic face processing within and outside of face-responsive brain regions

Abstract

Holistic processing is the tendency to perceive an object as an indecomposable whole, rather than by its parts. Psychological research has shown that faces are processed holistically and neuroimaging studies have linked holistic processing of faces to brain activity in face-responsive regions of the occipital-temporal cortex. However, as recent studies have suggested that other factors, such as Gestalt processing, may be involved in holistic face processing, we hypothesized that holistic face processing may not be unique to face-responsive brain regions. Using fMRI, we recorded the brain activity of human participants performing a composite face task. In this task, participants tend to judge same top face halves as different when they are aligned with different bottom face halves, as they are unable to ignore the irrelevant bottom face half information. We localized specific regions of interest defined by their responses to faces, scenes, objects and perceptual grouping, allowing us to investigate how activity in these regions changed during the composite face task. We found that the lateral occipital complex (LOC), fusiform face area 2 (FFA2) and transverse occipital sulcus (TOS) were sensitive to face alignment, suggesting a sensitivity of these regions to factors affecting holistic face processing. In addition, we found that the retrosplenial cortex (RSC) and the parahippocampal place area (PPA) showed a pattern of activity consistent with holistic processing of face identity, and the strength of this effect correlated with the strength of the behavioural composite effect measured with reaction times. These results suggest that holistic face processing occurs in brain regions involved in spatial and object processing, in addition to face-responsive brain regions, and that this neural activity directly relates to behavioural measures of holistic face processing.

Keywords: face perception, holistic processing, composite-face effect, fMRI

5.1. Introduction

Faces are perceived as indecomposable wholes, rather than by their separate component parts (e.g. eyes, nose, mouth), a phenomenon known as holistic processing (Farah et al., 1998; Maurer et al., 2002; Rossion, 2013). Holistic processing of faces has been demonstrated in psychological studies showing that people cannot selectively attend to one part of a face and ignore the rest of it (Maurer et al., 2002; Richler and Gauthier, 2014). For example, if the top-half of one person's face is aligned with the bottom-halves of two different faces (i.e. composite faces), observers have the tendency to perceive the two identical top-halves as two different identities, as they are unable to ignore the irrelevant bottom-halves of the faces. If the bottom-halves of the faces are misaligned from the top-halves, observers no longer process face holistically and they perceive the two top-halves to be the same. This phenomenon is known as the composite face effect (Hole, 1994; Young et al., 1987).

What neural processes underlie holistic processing of faces? Neuroimaging studies have identified neural activity consistent with holistic face processing in face-responsive regions of occipitotemporal cortex. Both the face-responsive occipital face area (OFA) and fusiform face area (FFA) have been shown to respond more to intact faces than to faces with the facial parts scrambled (Brandman and Yovel, 2016; Zhao et al., 2014a). Some studies have proposed that the FFA may process faces more holistically than the OFA. One study found higher responses to face parts than to whole faces in the OFA (Arcurio et al., 2012), and another study found that the FFA, but not the OFA, responds stronger when face parts are arranged in a normal configuration compared to a scrambled configuration (Liu et al., 2010). Behavioural studies have shown that inverted faces are processed less holistically than upright faces (Richler et al., 2011b; Rossion and Boremanse, 2008; Tanaka and Farah, 1993; Young et al., 1987), and correspondingly higher responses to upright as compared to inverted faces have been found in the FFA (Goffaux et al., 2013; Yovel and Kanwisher, 2005), but see also (Aguirre et al., 1999; Epstein et al., 2006; Grotheer et al., 2014; Haxby et al., 1999). Other neuroimaging studies have investigated which brain regions respond to the change in the perception of a face's identity when participants view composite faces. These studies found that changes in neural activity in the FFA, and in some cases also the OFA,

were consistent with changes in the perception of face identity induced by holistic processing (Andrews et al., 2010; Goffaux et al., 2013; Schiltz et al., 2010; Schiltz and Rossion, 2006). However, other studies found neural responses in the FFA consistent with a mixture of both holistic and part-based representations of faces (Harris and Aguirre, 2010, 2008). In combination, these studies show strong evidence that the FFA is involved in holistic face processing.

Many behavioural studies have demonstrated that holistic processing is not unique to faces. In particular, behavioural studies have demonstrated that objects of expertise can be processed holistically (Bukach et al., 2010; Diamond and Carey, 1986) and when participants are trained to recognise exemplars of novel kinds of objects this training leads to holistic processing of these objects (Chua and Gauthier, 2020; Gauthier and Tarr, 1997; Wong et al., 2009a). Neuroimaging studies investigating changes in neural processing related to expertise have shown that the FFA responds more to expertise objects in experts than in novices (Gauthier et al., 2000a; Xu, 2005), and have shown that the strength of these neural responses in the FFA correlates with behavioural measures of holistic processing of these objects (Gauthier and Tarr, 2002; Wong et al., 2009b). Furthermore, one study identified a correlation between the level of expertise and the amount of neural activity related to holistic processing of expertise objects in the anterior portion of the FFA, known as the FFA2 (Ross et al., 2018). These studies suggest that neural activity in the FFA may be involved in holistic processing of expertise objects.

These behavioural and neuroimaging studies provide strong evidence of a link between holistic processing, expertise and neural activity in the FFA. However, recent behavioural work suggests that other factors may also contribute to holistic face processing. One study demonstrated that non-expertise objects can be processed as holistically as faces, and that this may be linked to salient Gestalt information in these objects (Zhao et al., 2016). Two recent behavioural studies investigated if there is interference between holistic processing of faces and these non-expertise objects (Curby et al., 2019; Curby and Moerel, 2019). Based on their findings the authors proposed that holistic face processing may involve two components, an expertise component that overlaps with mechanisms relating to holistic processing of expertise objects, and a perceptual component that overlaps with holistic processing of non-expertise objects with salient Gestalt information (Curby and

Moerel, 2019). An open question is what neural mechanisms might underlie this perceptual, Gestalt-related component of holistic face processing.

In this study, we investigated if the neural mechanisms supporting holistic face processing involve broader brain networks beyond the face-responsive network. Most previous studies investigating holistic face processing specifically localized face-responsive brain regions, but did not localize other regions related to high-level object processing. Although activity in non-localized regions may be revealed using whole-brain analyses, it is well-known that activity may be masked due to suboptimal alignment of functional brain regions across participants and poor statistical power in these analyses (Saxe et al., 2006; Weiner and Grill-Spector, 2013). In the present study, we recorded brain activity using functional magnetic resonance imaging (fMRI) as participants performed a composite face task (Hole, 1994; Young et al., 1987). In this task, when the top-half of one person's face is aligned with the bottom-halves of two different faces (i.e. composite faces), observers have the tendency to perceive the two identical top-halves as two different identities. Participants viewed pairs of composite-faces and made same/different judgements as to the identity of the top-half of the face.

We localized a variety of regions of interest (ROIs) that are either face-responsive or are sensitive to information that may support holistic processing. For face-responsive ROIs, we localized the FFA and the OFA, which have been shown to be related to holistic processing of faces in previous studies (Andrews et al., 2010; Goffaux et al., 2013; Harris and Aguirre, 2010; Schiltz et al., 2010; Schiltz and Rossion, 2006). We subdivided the FFA into FFA1 and FFA2 (two components of the FFA) (Weiner et al., 2016, 2014), as some previous work has found evidence of holistic processing only in the FFA2 (Ross et al., 2018). Additionally, we localized a more recently defined, higher-level face-responsive brain region, the anterior temporal face area (ATFA) (Rajimehr et al., 2009; Tsao et al., 2008). Given that FFA has previously shown more consistent evidence of holistic processing than OFA, a lower-level region, we considered it possible that an even higher-level face processing region, ATFA, may also process faces holistically.

For ROIs outside of the face-responsive brain network, we first localized ROIs responsive to scenes, specifically scene-responsive transverse occipital sulcus (TOS, also referred to as occipital place area, OPA), parahippocampal place area (PPA) and

retrosplenial cortex (RSC). In the same way as the FFA responds more strongly to whole faces than facial parts, both PPA and RSC show higher neural activity for intact scenes than for fractured scenes (Kamps et al., 2016), suggesting that these areas are sensitive to configuration information. Although TOS is thought to be sensitive to the local elements of scenes (e.g. surfaces, furniture) (Kamps et al., 2016), it does contribute to the fine-grained perceptual discrimination of very similar scenes (Dilks et al., 2013) and it shows stronger activation to a whole scene (e.g. a furnished room) compared to scene components (e.g. isolated furniture) (Bettencourt and Xu, 2013). Furthermore, both PPA and TOS/OPA, as well as the FFA and the object-responsive lateral occipital complex (LOC), have been shown to have stronger responses to holistically processed scene stimuli as compared to control stimuli that were matched for low-level factors but not processed holistically (Schindler and Bartels, 2016). If configural/relational processing in general contributes to holistic face processing then we hypothesized that these scene-responsive areas may also exhibit neural activity related to holistic processing.

Secondly, we localized the object-responsive LOC, allowing us to test whether holistic face processing is a general mechanism of high-level visual object processing. Thirdly, we localized a region in the superior parietal lobule (SPL) that has been shown to be involved in holistic processing of highly controlled bi-stable Gestalt, occlusion and plaid stimuli (Grassi et al., 2018, 2016; Zaretskaya et al., 2013), and has also been found to be involved in processing of configural face information (Zachariou et al., 2017). As Gestalt information has been shown to be important for holistic processing (Zhao et al., 2016; Zhao and Bülthoff, 2017), we hypothesized that neural activity in this brain region might be involved in holistic processing of faces during the composite-face task.

5.2. Materials and methods

5.2.1. Participants

Nineteen participants (13 female, 6 male, 20-39 years old) were included in our fMRI data analyses. Data from three additional participants were excluded prior to the fMRI data analyses, one due to excessive head movement during scanning, two due to poor performance in the behavioural task (less than 65% correct responses on congruent-identity trials, where no illusion is present). A power analysis conducted using G*Power3 (Faul et al., 2007) indicated that a sample size of 19 is required to detect a medium effect size of η_p^2 = 0.06 at the 0.05 alpha level with 70% power. For correlation analyses, a power analysis indicated that a sample size of 19 is required to detect a large effect size (ρ = 0.5) at the 0.05 alpha level with 60% power. All participants provided written informed consent prior to the experiment, and the procedure was approved by the local ethics committee of the University Clinic Tübingen.

5.2.2. Stimuli

5.2.2.1. Main experiment stimuli

The experimental stimuli were created using images of 3D face models from the face database of the Max Planck Institute for Biological Cybernetics (Blanz and Vetter, 1999; Troje and Bülthoff, 1996). We selected the faces of 12 Caucasian individuals (6 females) and paired each face once with another face of the same sex to make 12 face pairs. Each face was separated into a top and bottom half, and the halves of the pairs were recombined to create composite faces, as illustrated by the 8 conditions in Fig. 1A. A horizontal black line (0.03° of visual angle) was shown between the top and bottom halves of each face to clearly separate the two face halves. During the experiment, face stimuli were displayed with a height of 3.9° and width of 3.0° of visual angle. For misaligned stimuli, the bottom half of the face was shifted 1.0° of visual angle to the left. Faces were grayscale, and were shown in front of a gray background. Stimuli used for the practice trials were created via the same method, using additional faces taken from the database.

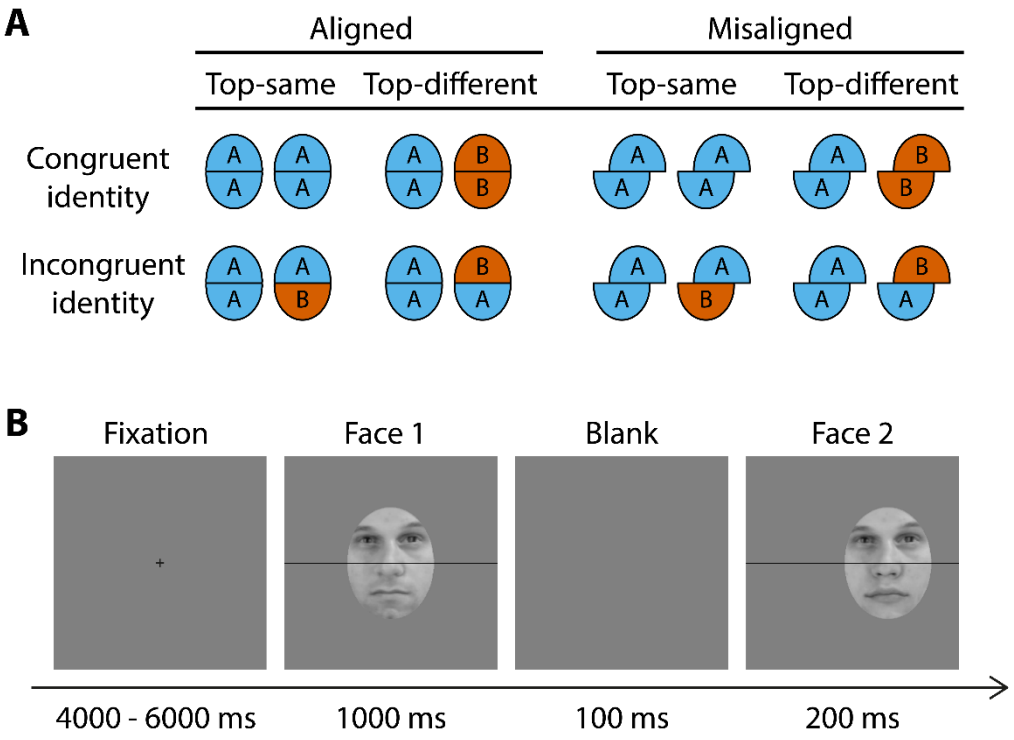

Figure 1. Experimental conditions and trial outline. (A) Experimental conditions. The conditions consisted of a 2 x 2 x 2 factorial design, with factors *alignment*, whether the top and bottom halves of the faces were aligned or misaligned, *top-same or top-different*, whether the top halves of the faces were the same or different from each other and *congruency*, whether the bottom half of face 2 was congruent with respect to the top half of face 2 or not (e.g. *congruent-identity* trials are when the bottom-half is the same if the top-half is the same and the bottom-half is different if the top-half is different). (B) Trial outline. Participants fixated for either 4 or 6 s, then viewed a first face, followed by a blank screen and then a second face. Participants then responded during the next fixation whether the top-halves of the two faces were the same or different.

5.2.2.2. Localizer stimuli

The localizer stimuli were grayscale images of faces, objects, scenes and phase-scrambled scenes (9 exemplars per category). Phase-scrambled scenes were Fourier-scrambled versions of the scene images.

5.2.3. Experimental design

Participants lay supine in the scanner and viewed the stimuli on a screen positioned behind their head, via a mirror attached to the head coil. The screen was positioned 82 cm from the participant, and spanned 28° x 16° of visual angle in horizontal and vertical directions respectively. Stimuli were presented via a projector with resolution 1920x1080. The experiment was programmed with Matlab 2013b using the Psychophysics Toolbox extensions (Brainard, 1997; Kleiner et al., 2007) on a Windows PC.

5.2.3.1. Main experiment procedure

Participants performed a composite face task while their brain activity was recorded using fMRI. On each trial participants viewed two faces and made a judgement whether the top-halves of the faces were the same or different. The experimental design consisted of 8 conditions of a 2 x 2 x 2 factorial design (see Fig. 1A). The factors were *alignment* (whether the bottom halves of the faces were aligned or misaligned with the top halves), *congruency*, (whether the bottom half of the second face was congruent with respect to the top half of the second face or not) and *top-same/top-different* (whether the top halves of the two faces were the same or different from each other). Each participant completed 3 runs, where each run contained 64 trials (8 repetitions per condition). Conditions were presented in a carryover counterbalanced design, such that each condition was preceded by every other condition once per run (Brooks, 2012). This was to avoid biases from carryover blood-oxygen-level dependent (BOLD) activation from a previous condition (Aguirre, 2007).

The trial procedure is illustrated in Fig. 1B. Participants viewed a central fixation cross for 4 s or 6 s (50% of trials each, order randomized). The first face was shown centrally on the screen for 1 s, followed by a blank screen (presented for 100 ms), then the second face was shown, 1.2° of visual angle offset to the right of the centre of the screen, for 200 ms. Participants responded using a button press whether they judged the top halves of the two faces to be the same of different. They were instructed to ignore the bottom halves of

the faces and to respond as quickly and accurately as possible. The fingers participants used to respond same/different were counterbalanced across participants.

Participants performed practice trials prior to the experiment to familiarise them with the task. Each participant performed 8 practice trials outside of the MRI scanner and 8 practice trials inside the MRI scanner.

5.2.3.2. Localizer experiment procedure

Participants completed 2 runs of the localizer experiment, which was used to define face-, scene- and object-responsive ROIs. In each run, participants viewed blocks containing faces, scenes, objects and phase-scrambled scenes. Faces and objects were shown in front of the phase-scrambled scenes to keep the visual field size of the stimuli constant in all blocks (scene images were equal in size to the phase-scrambled scenes). Blocks were presented in a carryover counterbalanced sequence (Brooks, 2012). In each block 8 images were shown, where each image was shown for 1.8 s, followed by a 0.2 s blank, grey screen. Participants performed a one-back task on the images (repetitions once every 9 s on average) to keep their attention to the stimuli.

5.2.4. Imaging parameters

Images were acquired using a 3T Siemens Prisma scanner with a 64-channel head coil (Siemens, Erlangen, Germany). Functional T2* echoplanar images (EPI) were acquired using a sequence with the following parameters; multiband acceleration factor 2, TR 1.39 s, TE 30 ms, flip angle 68°, FOV 192x192 mm. Volumes consisted of 42 slices, with an isotropic voxel size of 3x3x3 mm. The first 8 volumes of each run were discarded to allow for equilibration of the T1 signal. For each participant a high-resolution T1-weighted anatomical scan was acquired with the following parameters; TR 2 s, TE 3.06 ms, FOV 232x256 mm, 192 slices, isotropic voxel size of 1x1x1 mm.

5.2.5. fMRI data preprocessing

fMRI data was preprocessed with SPM12 (http://www.fil.ion.ucl.ac.uk/spm/). Functional images were slice-time corrected, realigned and coregistered to the anatomical image. The images were then normalized to MNI (Montreal Neurological Institute) space and spatially smoothed with a 6 mm full-width at half-maximum Gaussian kernel.

5.2.6. Definition of regions of interest

Figure 2 illustrates the average locations of our regions of interest (ROIs) and Table 1 shows the mean MNI coordinates and volumes of each ROI. We defined face-, scene- and object-responsive ROIs using data from the localizer runs. Firstly, the contrast faces > objects and scenes was used to define the OFA, FFA1, FFA2 and ATFA (Gauthier et al., 2000b; Kanwisher et al., 1997; Rajimehr et al., 2009; Tsao et al., 2008). We defined the FFA1 and FFA2 based on functional selectivity and previously described anatomical landmarks (Weiner et al., 2016, 2014). Secondly, the contrast scenes > faces and objects was used to define the TOS, RSC and PPA (Epstein and Kanwisher, 1998; Grill-Spector, 2003; Maguire, 2001). Thirdly, the contrast objects > phase-scrambled scenes was used to define the LOC (Malach et al., 1995). We defined each ROI individually in each participant, by selecting all active voxels falling within spheres (radius 6 mm) centred on the peak of activity in each hemisphere. A threshold of $p < 0.001$ uncorrected was used to define active voxels.

We additionally defined SPL and V1, based on anatomical location and higher activity during stimulus presentation (including all conditions) compared to the fixation interval between trials. This contrast is orthogonal to the activity differences between the conditions in this study (Friston et al., 2006). We used a $p < 0.05$ familywise error rate (FWE) corrected threshold to define voxels more active during the stimulus than fixation. SPL was defined by selecting all active voxels falling within spheres (radius 6 mm) centred on the peak of activity in superior parietal cortex of each hemisphere. The entire V1 was defined first for each participant using anatomical labels generated by Freesurfer (Hinds et al., 2009) (https://surfer.nmr.mgh.harvard.edu/). To define our final V1 ROI, we selected all posterior V1 voxels that were more active when participants viewed the pairs of face stimuli as compared to when they fixated and viewed a grey screen. Participants could move their eyes when viewing the faces, therefore this V1 ROI reflects the V1 voxels activated for each individual participant when viewing the face stimuli.

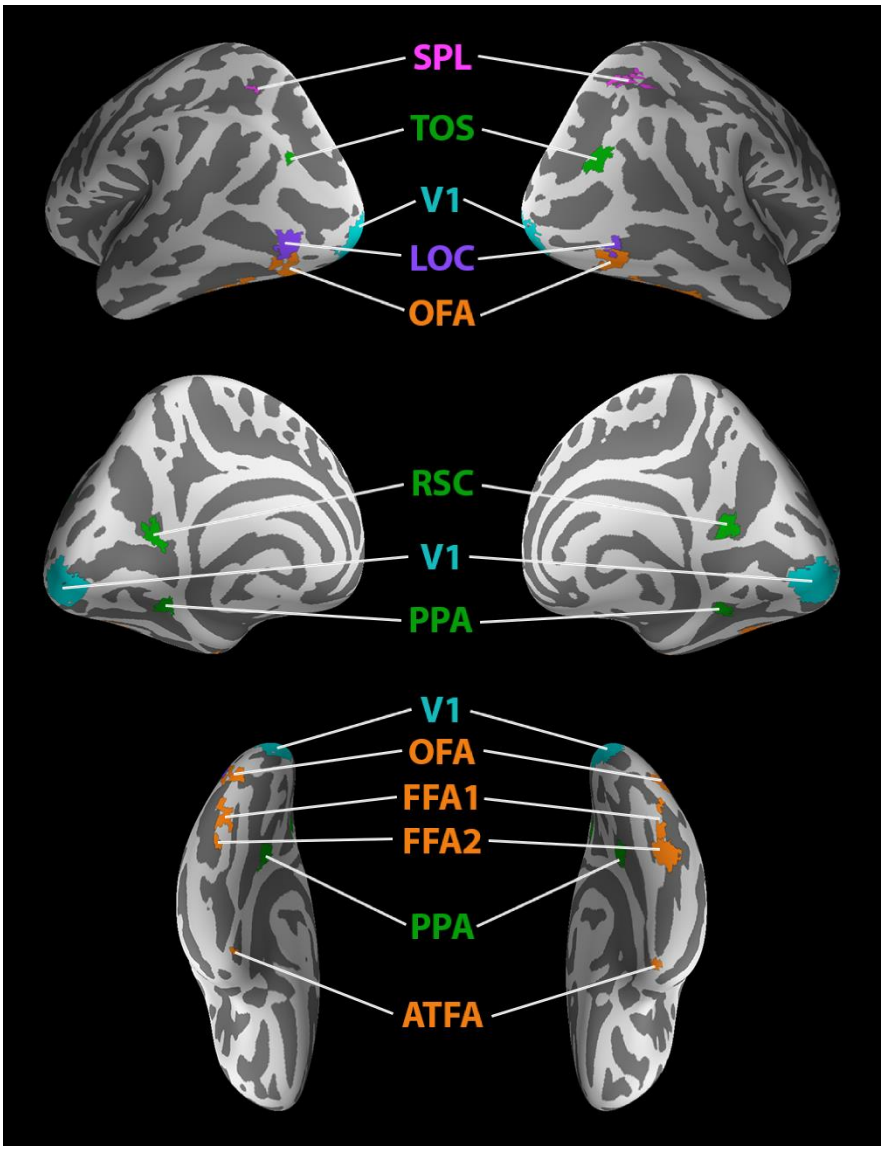

Figure 2. Locations of ROIs. ROIs include face-responsive OFA (occipital face area), FFA1 (fusiform face area 1), FFA2 (fusiform face area 2) and ATFA (anterior temporal face area) shown in orange, scene-responsive TOS (transverse occipital sulcus), RSC (retrosplenial cortex) and PPA (parahippocampal place area) shown in green, object-responsive LOC (lateral occipital complex) shown in purple, parietal SPL (superior parietal lobule) shown in magenta and V1 shown in cyan. ROIs were defined individually in volume-space for each participant, for visualisation in this figure we show group average ROIs projected onto the inflated cortical surface. We defined group ROIs using a relatively low threshold as some information was lost during projection to the cortical surface. Thus, voxels were included in each group average ROI if they were part of the ROI in at least 25% of participants.

Table 1. ROI locations and volumes

Average x, y and z coordinates (in MNI space) and volume of each ROI (± standard deviations). N indicates the number of participants each ROI was identified in.

ROI	hem	x	y	z	Volume (mm^3)	N
OFA	left	-39 ± 4.6	-81 ± 4.3	-10 ± 4.0	197 ± 42.4	19
	right	42 ± 4.1	-79 ± 4.3	-10 ± 4.0	208 ± 34.3	19
FFA1	left	-40 ± 4.1	-62 ± 8.8	-17 ± 4.0	202 ± 35.6	18
	right	42 ± 5.1	-63 ± 6.6	-16 ± 4.1	204 ± 44.8	18
FFA2	left	-41 ± 4.0	-43 ± 10.2	-21 ± 5.3	157 ± 61.4	16
	right	42 ± 3.4	-44 ± 5.5	-19 ± 3.6	205 ± 36.5	17
ATFA	left	-35 ± 5.3	-8 ± 5.2	-34 ± 6.0	86 ± 67.4	14
	right	34 ± 3.6	-5 ± 4.9	-38 ± 4.8	134 ± 60.7	16
TOS	left	-32 ± 5.7	-85 ± 4.9	22 ± 7.0	203 ± 35.5	19
	right	37 ± 4.1	-80 ± 2.9	21 ± 7.4	213 ± 22.1	19
RSC	left	-17 ± 3.2	-59 ± 3.3	14 ± 3.6	184 ± 48.3	18
	right	19 ± 3.0	-57 ± 4.8	17 ± 4.9	198 ± 54.1	18
PPA	left	-26 ± 2.8	-44 ± 3.9	-10 ± 3.4	197 ± 52.9	19
	right	29 ± 2.6	-45 ± 5.5	-10 ± 3.3	212 ± 19.4	19
LOC	left	-43 ± 4.1	-79 ± 4.3	-4 ± 4.6	222 ± 13.6	19
	right	43 ± 3.8	-80 ± 5.1	-4 ± 6.0	214 ± 25.6	19
SPL	left	-26 ± 4.7	-60 ± 6.8	50 ± 6.1	206 ± 49.2	19
	right	28 ± 4.9	-56 ± 6.7	50 ± 5.9	213 ± 49.8	19
V1	left	-13 ± 3.2	-98 ± 2.7	-7 ± 4.0	911 ± 380.6	19
	right	13 ± 3.2	-96 ± 1.6	-4 ± 4.1	911 ± 211.5	19

5.2.7. Statistical Analyses

5.2.7.1. Behavioural analyses

Participants were instructed to respond whether they judged the top-halves of the face pairs to be the same or different. We calculated our participants' behavioural performance with accuracy (% correct) and reaction times. For each behavioural measure we first performed a 2 (alignment) x 2 (congruency) x 2 (top-same/top-different) repeated measures ANOVA to investigate whether there was a significant three-way interaction between the three factors. We then separated the top-same and top-different conditions to investigate if we would find behavioural responses consistent with holistic processing. We separated the top-same and top-different conditions due to previous evidence that evidence of holistic processing is stronger for top-same than top-different conditions (Goffaux, 2012; Goffaux et al., 2013). For both top-same and top-different conditions we performed 2 (alignment) x 2 (congruency) repeated measures ANOVAs to test for an interaction effect between congruency and alignment and/or an effect of congruency. Due to holistic processing, we expected to find a difference in behavioural performance between the congruent-identity and incongruent-identity conditions that was larger for the aligned conditions compared to the misaligned conditions. The rationale is the following. For accuracy, a lower performance is expected in aligned incongruent-identity conditions compared to aligned congruent-identity conditions, due to participants being unable to ignore the irrelevant bottom face half information. This difference should be reduced in the misaligned conditions, as here participants are able to ignore this bottom face half information. For reaction times, a longer reaction time is expected for aligned incongruent-identity conditions compared to aligned congruent-identity conditions, due to participants taking longer to make their decision for this condition. Again, this difference should be reduced in the misaligned conditions as participants are able to ignore the irrelevant bottom face half information. In cases where we found significant effects in our ANOVA results, we performed follow-up *t*-tests to confirm that the pattern of behavioural performance matched these expectations.

5.2.7.2. fMRI analyses

We modelled a GLM for each participant containing regressors for our 8 conditions, plus 6 realignment regressors from the motion correction, using SPM12. The 8 condition

regressors modelled the activity to each trial of the condition, excluding any trials where the participant did not make a task response (i.e. did not press a button to indicate a same/different judgment of the faces, on average 1.8% of trials). Responses to each condition are reported in % signal change with respect to the baseline of the GLM in each ROI. We performed a 2 (alignment) x 2 (congruency) x 2 (top-same/top-different) repeated measures ANOVA to investigate whether any brain regions showed a triple interaction effect between the 3 factors in neural activity. We then performed further analyses investigating specific aspects of composite-face effect.

We investigated the effect of alignment, as aligned faces are considered to be processed more holistically than misaligned ones (Rossion, 2013; Young et al., 1987), therefore we considered that regions involved in holistic processing of faces would show differences in activity between these conditions. We included only congruent-identity conditions in this analysis, as the perception of face identity differs between aligned and misaligned incongruent conditions. Therefore differences in neural activation between aligned and misaligned incongruent conditions could reflect differences in neural activity related to the perception of face identity, rather than related to holistic processing. We performed 2 (alignment) x 2 (top-same/top-different) repeated measures ANOVAs and investigated if any regions showed an effect of alignment.

Next, we investigated whether any brain regions would show a difference in neural responses between congruent-identity and incongruent-identity conditions that was larger for the aligned conditions compared to the misaligned conditions. We predicted that any brain regions encoding face identity in a holistic manner would show this pattern of responses, due to the integration of top- and bottom-half face information for the aligned conditions, but not the misaligned ones. In contrast, we predicted that any brain regions encoding face identity in a parts-based manner would respond similarly to aligned and misaligned faces, and thus not show an interaction between congruency and alignment. To test for this specific pattern of responses, we used 2 (alignment) x 2 (congruency) repeated measures ANOVAs to investigate if any brain regions show an interaction between congruency and alignment and/or an effect of congruency. We then performed follow-up *t*-tests in any regions showing significant effects to investigate if this was due to a difference in neural activity between the congruent-identity and incongruent-identity conditions that

was larger for the aligned conditions compared to the misaligned conditions. We performed this analysis separately for top-same and top-different conditions due to the known differences in the strength of holistic processing effects between these conditions (Goffaux, 2012; Goffaux et al., 2013).

We performed all fMRI analyses in our ten ROIs as well as in whole-brain analyses. For ROI analyses, we corrected for multiple comparisons using a Bonferroni-correction to adjust for the number of ROIs tested. For whole-brain analyses, we used a False Discovery Rate (FDR) correction to adjust for multiple comparisons.

5.2.7.3. Analyses linking behaviour and neural activity

We performed follow-up analyses comparing any effects identified related to holistic processing in neural and behavioural responses. We performed Pearson's correlation analyses between accuracy and reaction time results in behaviour, and neural responses in ROIs showing significant effects related to holistic processing, in order to assess whether there was a link between the strength of these effects for each participant in behavioural responses and neural activity. We corrected for multiple comparisons using a Bonferroni-correction to adjust for the number of ROIs tested.

5.2.8. Data and code availability statement

Data cannot be shared as participants were informed that their data would be stored confidentially, in accordance with the rules of the local ethics committee. Code is available on request.

5.3. Results

5.3.1. Behavioural results

We measured participants' behavioural performance in the composite-face task during scanning using accuracy (% correct) and reaction times. For both behavioural measures 2 x 2 x 2 repeated-measures ANOVAs showed a significant triple interaction between alignment, congruency and top-same/top-different conditions (accuracy: $F_{1,18}$ = 24.24, p = 1.1 × 10^{-4}, η_p^2 = 0.57; reaction times: $F_{1,18}$ = 8.92, p = 0.0079, η_p^2 = 0.33). We conducted further behavioural analyses separately for top-same and top-different conditions as there are known differences in the behavioural responses to these conditions (Goffaux, 2012; Goffaux et al., 2013).

5.3.1.1. Accuracy and reaction times for top-same conditions

Accuracy (% correct) and reaction times for the top-same conditions are shown in Figure 3. For both measures, 2 (alignment) x 2 (congruency) repeated-measures ANOVAs revealed both a significant effect of congruency (accuracy: $F_{1,18}$ = 23.67, p = 1.2 × 10^{-4}, η_p^2 = 0.57; reaction times: $F_{1,18}$ = 15.08, p = 0.0011, η_p^2 = 0.46) and a significant interaction between congruency and alignment (accuracy: $F_{1,18}$ = 32.12, p = 2.2 × 10^{-5}, η_p^2 = 0.64; reaction times: $F_{1,18}$ = 23.44, p = 1.3 × 10^{-4}, η_p^2 = 0.57). Furthermore, paired-sample t-tests showed that the congruency effect was significant for the aligned conditions (accuracy: M = 28.07 %, SE = 5.04 %; t_{18} = 5.57, p = 2.7 × 10^{-5}, Cohen's d_z = 1.28; reaction times: M = 0.13 s, SE = 0.023 s; t_{18} = 5.51, p = 3.1 × 10^{-5}, Cohen's d_z = 1.26) but not for the misaligned conditions (accuracy: M = 1.97 %, SE = 2.07 %; t_{18} = 0.95, p = 0.35, Cohen's d_z = 0.22; reaction times: M = 0.019 s, SE = 0.020 s; t_{18} = 0.92, p = 0.37, Cohen's d_z = 0.21). These results show that there is evidence of holistic processing elicited by the top-same conditions in our composite-task.

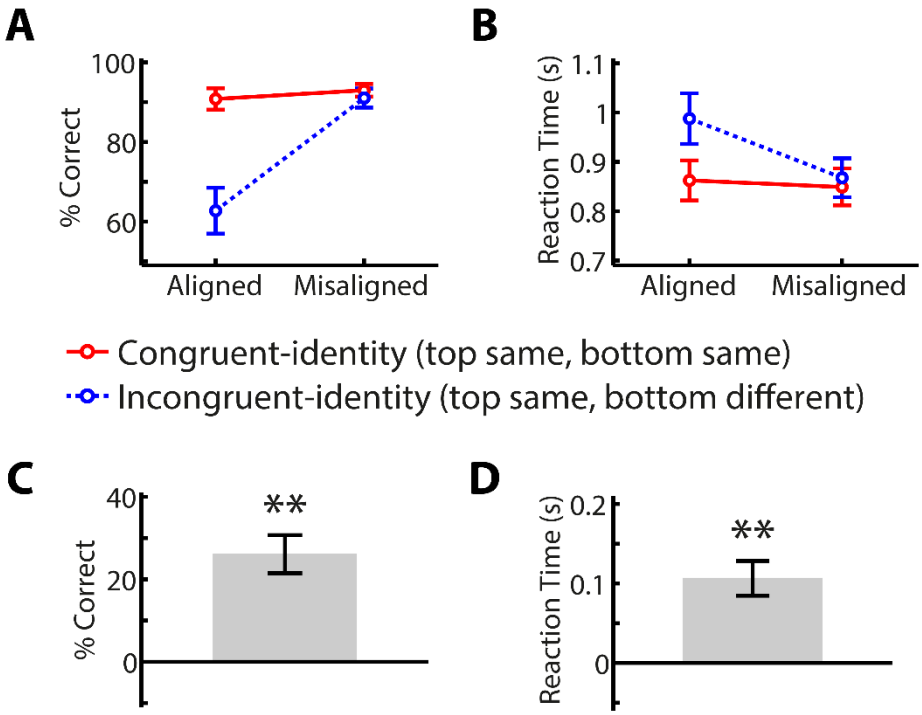

Figure 3. Behavioural performance in the top-same conditions of the composite-face task. (A) shows accuracy (% correct) as a function of congruency and alignment and (B) shows reaction times as a function of congruency and alignment. (C) and (D) show the interaction effect between congruency and alignment (difference between aligned congruent-identity and incongruent-identity conditions, minus the difference between misaligned congruent-identity and incongruent-identity conditions) as measured with accuracy (C) and reaction times (D). Error bars indicate ±1 *SEM*. ** indicates $p < 0.001$.

5.3.1.2. Accuracy and reaction times for top-different conditions

Accuracy (% correct) and reaction times for the top-different conditions are shown in Figure 4. In contrast to the top-same condition results, 2 (alignment) x 2 (congruency) repeated-measures ANOVAs showed no significant effect of congruency (accuracy: $F_{1,18} = 1.35$, $p = 0.26$, $\eta_p^2 = 0.070$; reaction times: $F_{1,18} = 0.25$, $p = 0.63$, $\eta_p^2 = 0.014$) or significant interaction between congruency and alignment (accuracy: $F_{1,18} = 0.39$, $p = 0.54$, $\eta_p^2 = 0.021$; reaction times: $F_{1,18} = 1.03$, $p = 0.32$, $\eta_p^2 = 0.054$). Thus, in our composite-task the holistic processing effect, as expected, seems to be driven by the top-same conditions rather than the top-different conditions.

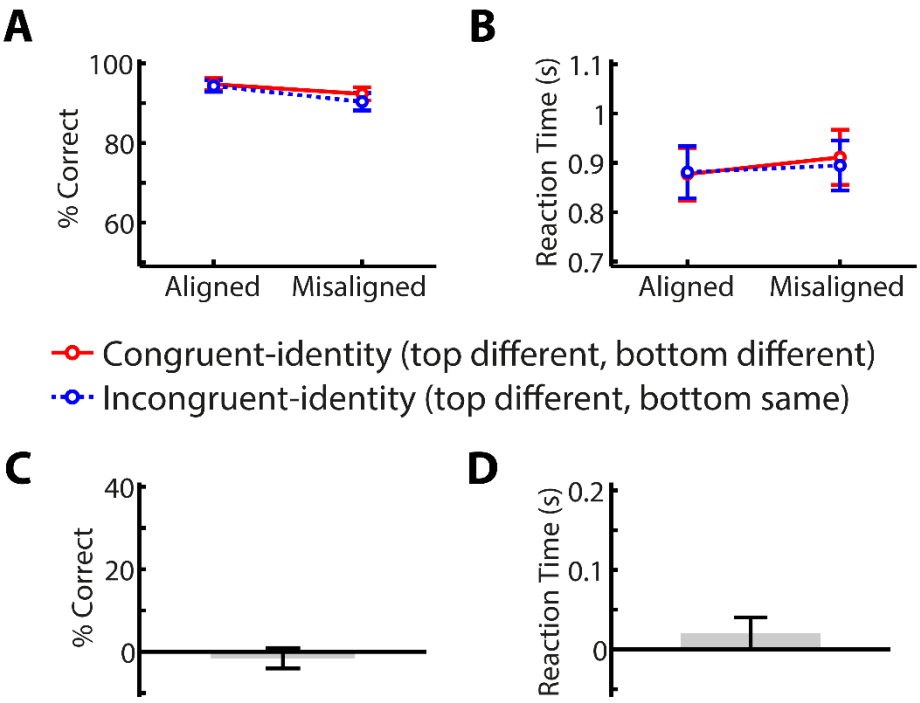

Figure 4. Behavioural performance in the top-different conditions of the composite-face task. (A) shows accuracy (% correct) as a function of congruency and alignment and (B) shows reaction times as a function of congruency and alignment. (C) and (D) show the interaction effect between congruency and alignment (difference between aligned congruent-identity and incongruent-identity conditions, minus the difference between misaligned congruent-identity and incongruent-identity conditions) as measured with accuracy (C) and reaction times (D). Error bars indicate ±1 *SEM*.

5.3.2. fMRI results

We first tested whether any brain regions showed a triple interaction between alignment, congruency and top-same/top-different conditions. 2 x 2 x 2 repeated measures ANOVAs showed a marginal triple interaction in the RSC ($F_{1,17}$ = 8.66, p = 0.0091, η_p^2 = 0.34) which did not survive Bonferroni-correction for N = 10 ROIs. None of the other ROIs tested showed a significant triple interaction effect (OFA: $F_{1,18}$ = 0.55, p = 0.47, η_p^2 = 0.030; FFA1: $F_{1,17}$ = 0.46, p = 0.51, η_p^2 = 0.027; FFA2: $F_{1,16}$ = 0.25, p = 0.62, η_p^2 = 0.015; ATFA: $F_{1,16}$ = 2.31, p = 0.15, η_p^2 = 0.13; TOS: $F_{1,18}$ = 0.85, p = 0.37, η_p^2 = 0.045; PPA: $F_{1,18}$ = 2.63, p = 0.12, η_p^2 = 0.13; LOC: $F_{1,18}$ = 0.86, p = 0.37, η_p^2 = 0.046; SPL: $F_{1,18}$ = 0.41, p = 0.53, η_p^2 = 0.022; V1: $F_{1,18}$ = 0.12, p = 0.73, η_p^2 = 0.0065).

We further tested for differences between specific contrasts of interest in order to investigate different aspects of the composite-face effect. We first investigated which regions show differences in neural activity based on how holistically the stimuli is processed by comparing neural responses to aligned and misaligned faces. In this analysis we included only the congruent-identity conditions (i.e. where top and bottom halves of the face pairs are both the same or both different) as in these conditions there is no change in the perception of the identity of the face between the aligned and misaligned conditions.

Secondly, we investigated which regions show differences in neural activity consistent with processing face identity in a holistic manner. To do this, we investigated which regions show a larger difference in neural activity between aligned congruent-identity and incongruent-identity face pairs than between misaligned congruent-identity and incongruent-identity face pairs. We conducted this analysis separately for top-same and top-different conditions due to the differences in participants' behavioural responses during the top-same and top-different conditions (see Figures 3-4).

5.3.2.1. Neural responses to alignment

We tested for differences in neural activity between aligned and misaligned congruent-identity conditions by investigating which regions showed a main effect of alignment in a 2 (alignment) x 2 (top-same/top-different) repeated measures ANOVA. For face-responsive ROIs (Fig. 5A), we found a marginally significant effect of alignment in the FFA2 ($F_{1,16}$ = 6.72, p = 0.020, η_p^2 = 0.30), which did not survive Bonferroni-correction for N =

10 ROIs. This marginal effect was driven by higher responses to aligned faces than misaligned ones (M = 0.31 %, SE = 0.12 %). We found no difference in responses to aligned and misaligned faces in any of the other face-responsive ROIs we tested (OFA: $F_{1,18}$ = 0.60, p = 0.45, η_p^2 = 0.032; FFA1: $F_{1,17}$ = 0.071, p = 0.79, η_p^2 = 0.0042; ATFA: $F_{1,16}$ = 2.33, p = 0.15, η_p^2 = 0.13).

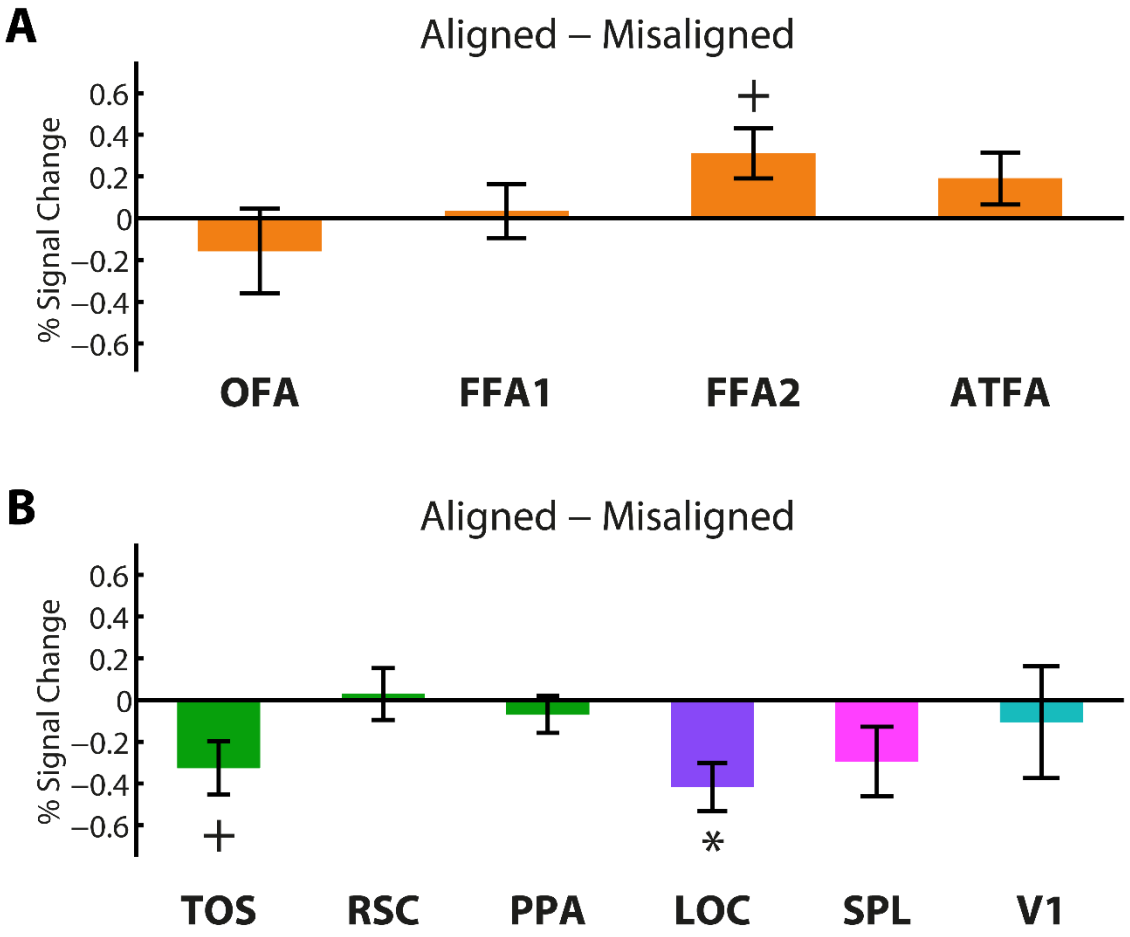

Figure 5. Differences in neural responses to aligned and misaligned faces for the congruent-identity conditions (i.e. where top and bottom halves of the face pairs were both the same or both different). We used the contrast aligned minus misaligned faces to investigate differences in neural activity in face-responsive ROIs (A) and in all other ROIs tested (B). Error bars indicate ±1 *SEM*. * indicates p < 0.05 Bonferroni-corrected for N = 10 ROIs, + indicates p < 0.05 uncorrected.

We further tested whether other ROIs (scene-responsive, object-responsive, parietal and early-visual ROIs) would show differences in neural activity between aligned and misaligned faces (Fig. 5B). We found a significant effect of alignment in the LOC ($F_{1,18}$ = 12.85, p = 0.0021, η_p^2 = 0.42), which survived Bonferroni-correction for N = 10 ROIs, and a marginally significant effect of alignment in the TOS ($F_{1,18}$ = 6.43, p = 0.021, η_p^2 = 0.26), which did not survive Bonferroni-correction for N = 10 ROIs. These effects were driven by higher responses to misaligned faces compared to aligned ones (LOC: M = 0.42 %, SE = 0.12 %; TOS: M = 0.32 %, SE = 0.13 %). None of the other ROIs tested showed significant differences in responses between aligned and misaligned faces (RSC: $F_{1,17}$ = 0.055, p = 0.82, η_p^2 = 0.0032; PPA: $F_{1,18}$ = 0.59, p = 0.45, η_p^2 = 0.032; SPL: $F_{1,18}$ = 3.14, p = 0.093, η_p^2 = 0.15; V1: $F_{1,18}$ = 0.16, p = 0.70, η_p^2 = 0.0086). We additionally performed a whole-brain analysis to investigate if any other regions showed differences in activity between the aligned and misaligned congruent-identity conditions. We did not identify any regions in this whole-brain analysis.

5.3.2.2. Interaction between congruency and alignment for top-same conditions

We tested whether any regions showed a difference in the neural responses to the aligned congruent-identity and incongruent-identity conditions that was reduced between the misaligned congruent-identity and incongruent-identity conditions. We first tested these differences for the top-same conditions, where we also found significant differences in participants' pattern of behavioural responses to these conditions (Section 5.3.1.1.).

For the face-responsive brain regions (Fig. 6), 2 (congruency) x 2 (alignment) repeated measures ANOVAs showed no significant interaction between congruency and alignment (OFA: $F_{1,18}$ = 3.07, p = 0.097, η_p^2 = 0.15; FFA1: $F_{1,17}$ = 0.79, p = 0.39, η_p^2 = 0.044; FFA2: $F_{1,16}$ = 0.91, p = 0.35, η_p^2 = 0.054; ATFA: $F_{1,16}$ = 3.35, p = 0.086, η_p^2 = 0.17) or significant effect of congruency (OFA: $F_{1,18}$ = 0.35, p = 0.56, η_p^2 = 0.019; FFA1: $F_{1,17}$ = 2.54, p = 0.13, η_p^2 = 0.13; FFA2: $F_{1,16}$ = 0.23, p = 0.63, η_p^2 = 0.015; ATFA: $F_{1,16}$ = 0.40, p = 0.53, η_p^2 = 0.025) in any of the face-responsive ROIs tested.

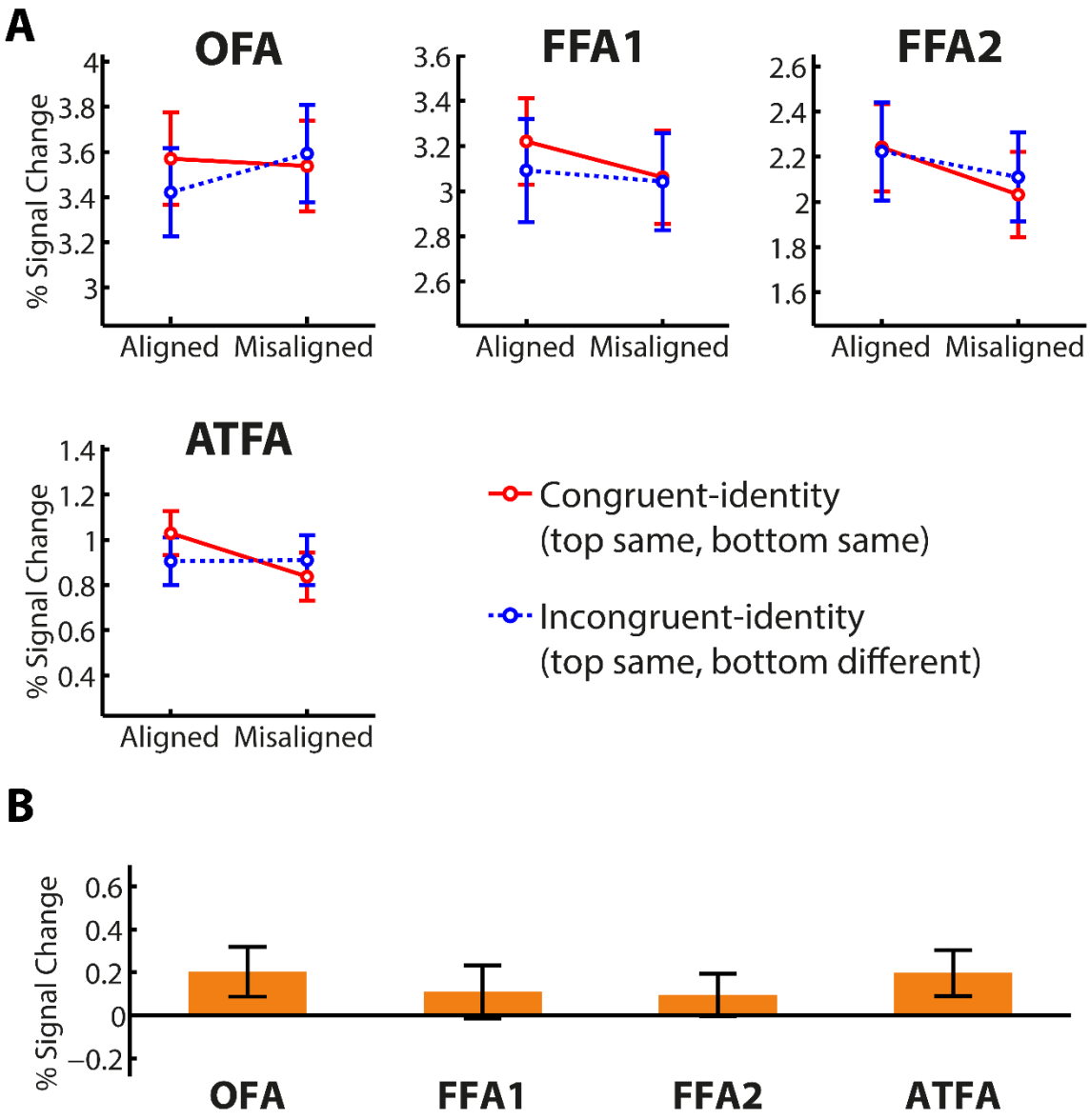

Figure 6. Neural responses to the top-same conditions in the face-responsive ROIs. (A) % signal change as a function of congruency and alignment. (B) Interaction between congruency and alignment with the contrast: *(aligned congruent-identity – aligned incongruent-identity) – (misaligned congruent-identity – misaligned incongruent-identity)*. Error bars indicate ±1 *SEM*.

We further tested whether other ROIs, including scene-responsive, object-responsive, parietal and early-visual ROIs, would show an interaction between congruency and alignment or an effect of congruency (Fig. 7). For the scene responsive ROIs, we found a significant interaction between congruency and alignment in the RSC ($F_{1,17}$ = 14.07, p = 0.0016, η_p^2 = 0.45) and the PPA ($F_{1,18}$ = 11.58, p = 0.0032, η_p^2 = 0.39), both surviving Bonferroni-correction for N = 10 ROIs. We did not find a significant interaction between congruency and alignment in the scene-responsive TOS ($F_{1,18}$ = 1.85, p = 0.19, η_p^2 = 0.093) and we did not find a significant effect of congruency in any of the scene-responsive ROIs (RSC: $F_{1,17}$ = 3.55, p = 0.077, η_p^2 = 0.17; PPA: $F_{1,18}$ = 4.33, p = 0.052, η_p^2 = 0.19; TOS: $F_{1,18}$ = 3.98, p = 0.061, η_p^2 = 0.18). We performed follow-up paired t-tests in the RSC and PPA to confirm that the interaction effect was due to a larger difference between congruent-identity and incongruent-identity aligned conditions than between congruent-identity and incongruent-identity misaligned conditions, the expected pattern for an effect induced by holistic processing. We found a significant effect of congruency for the aligned conditions in both RSC (M = 0.32 %, SE = 0.082 %; t_{17} = 3.91, p = 0.0011, Cohen's d_z = 0.92) and PPA (M = 0.24 %, SE = 0.061 %; t_{18} = 3.87, p = 0.0011, Cohen's d_z = 0.89), but not for the misaligned conditions in either ROI; RSC (M = -0.058 %, SE = 0.090 %; t_{17} = -0.64, p = 0.53, Cohen's d_z = -0.15), PPA (M = -0.018 %, SE = 0.067 %; t_{18} = -0.27, p = 0.79, Cohen's d_z = -0.063). Interestingly, the positive effect direction shows that the effect we identified was due to a repetition enhancement effect, i.e. higher activity in aligned congruent-identity conditions compared to aligned incongruent-identity conditions (see Section 5.4.2. for a discussion of the repetition-effect direction).

None of the other ROIs we tested showed a significant interaction between congruency and alignment (LOC: $F_{1,18}$ = 2.23, p = 0.15, η_p^2 = 0.11; SPL: $F_{1,18}$ = 0.095, p = 0.76, η_p^2 = 0.0052; V1: $F_{1,18}$ = 1.09, p = 0.31, η_p^2 = 0.057) or a significant effect of congruency (LOC: $F_{1,18}$ = 1.91, p = 0.18, η_p^2 = 0.096; SPL: $F_{1,18}$ = 1.42, p = 0.25, η_p^2 = 0.073; V1: $F_{1,18}$ = 0.52, p = 0.48, η_p^2 = 0.028). We additionally performed a whole-brain analysis to look if any other regions showed an interaction between congruency and alignment or effect of congruency, but we did not identify any regions in this analysis.

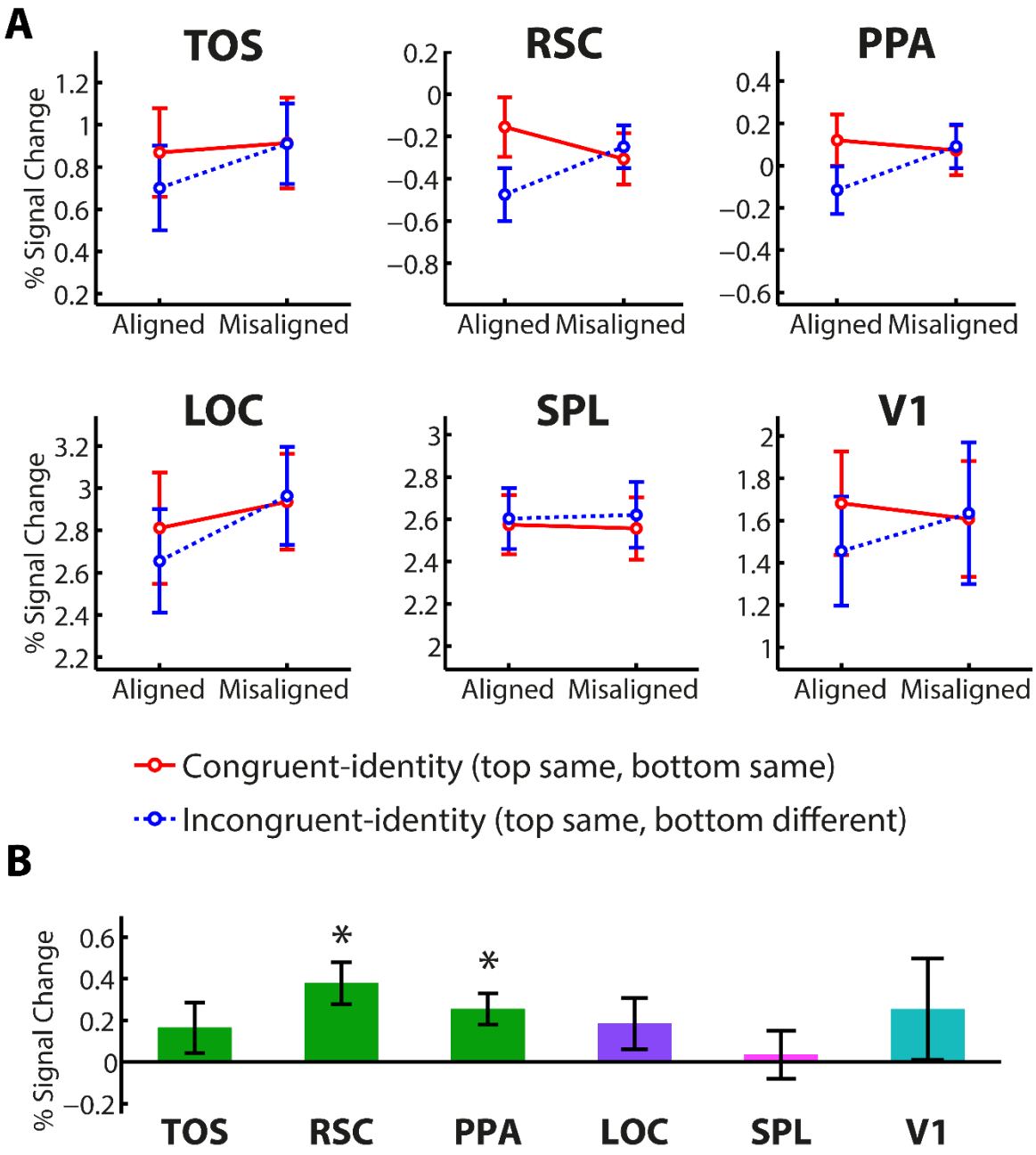

Figure 7. Neural responses to the top-same conditions in the scene-responsive, object-responsive, perceptual grouping and early visual ROIs. (A) % signal change as a function of congruency and alignment. (B) Interaction between congruency and alignment with the contrast: *(aligned congruent-identity – aligned incongruent-identity) – (misaligned congruent-identity – misaligned incongruent-identity)*. Error bars indicate ±1 *SEM*. * indicates $p < 0.05$, Bonferroni-corrected for N = 10 ROIs.

5.3.2.3. Interaction between congruency and alignment for top-different conditions

We tested whether any regions showed an interaction between congruency and alignment consistent with holistic processing for the top-different conditions (Fig. 8 & 9). The impact of these conditions on the behavioural measurement of holistic processing has been debated (Richler and Gauthier, 2013; Rossion, 2013). We aimed to investigate if they have an impact on neural activity.

Consistent with our behavioural results, none of our ROIs showed a significant interaction between congruency and alignment (OFA: $F_{1,18}$ = 5.11 x 10^{-4}, p = 0.98, η_p^2 = 2.8 x 10^{-5}; FFA1: $F_{1,17}$ = 0.11, p = 0.74, η_p^2 = 0.0065; FFA2: $F_{1,16}$ = 0.022, p = 0.88, η_p^2 = 0.0014; ATFA: $F_{1,16}$ = 0.44, p = 0.52, η_p^2 = 0.027; TOS: $F_{1,18}$ = 0.22, p = 0.65, η_p^2 = 0012; RSC: $F_{1,17}$ = 0.79, p = 0.39, η_p^2 = 0.044; PPA: $F_{1,18}$ = 0.031, p = 0.86, η_p^2 = 0.0017; LOC: $F_{1,18}$ = 0.089, p = 0.77, η_p^2 = 0.0049; SPL: $F_{1,18}$ = 0.51, p = 0.48, η_p^2 = 0.028; V1: $F_{1,18}$ = 0.065, p = 0.80, η_p^2 = 0.0036) or a significant effect of congruency (OFA: $F_{1,18}$ = 2.05, p = 0.17, η_p^2 = 0.10; FFA1: $F_{1,17}$ = 0.11, p = 0.74, η_p^2 = 0.0064; FFA2: $F_{1,16}$ = 0.18, p = 0.68, η_p^2 = 0.011; ATFA: $F_{1,16}$ = 4.06, p = 0.061, η_p^2 = 0.20; TOS: $F_{1,18}$ = 0.42, p = 0.52, η_p^2 = 0.023; RSC: $F_{1,17}$ = 0.44, p = 0.52, η_p^2 = 0.025; PPA: $F_{1,18}$ = 0.19, p = 0.67, η_p^2 = 0.010; LOC: $F_{1,18}$ = 2.77, p = 0.11, η_p^2 = 0.13; SPL: $F_{1,18}$ = 0.60, p = 0.45, η_p^2 = 0.032; V1: $F_{1,18}$ = 0.029, p = 0.87, η_p^2 = 0.0016). We additionally performed a whole-brain analysis to investigate if any other regions showed an interaction between congruency and alignment or an effect of congruency during the top-different conditions. No regions were identified in this analysis.

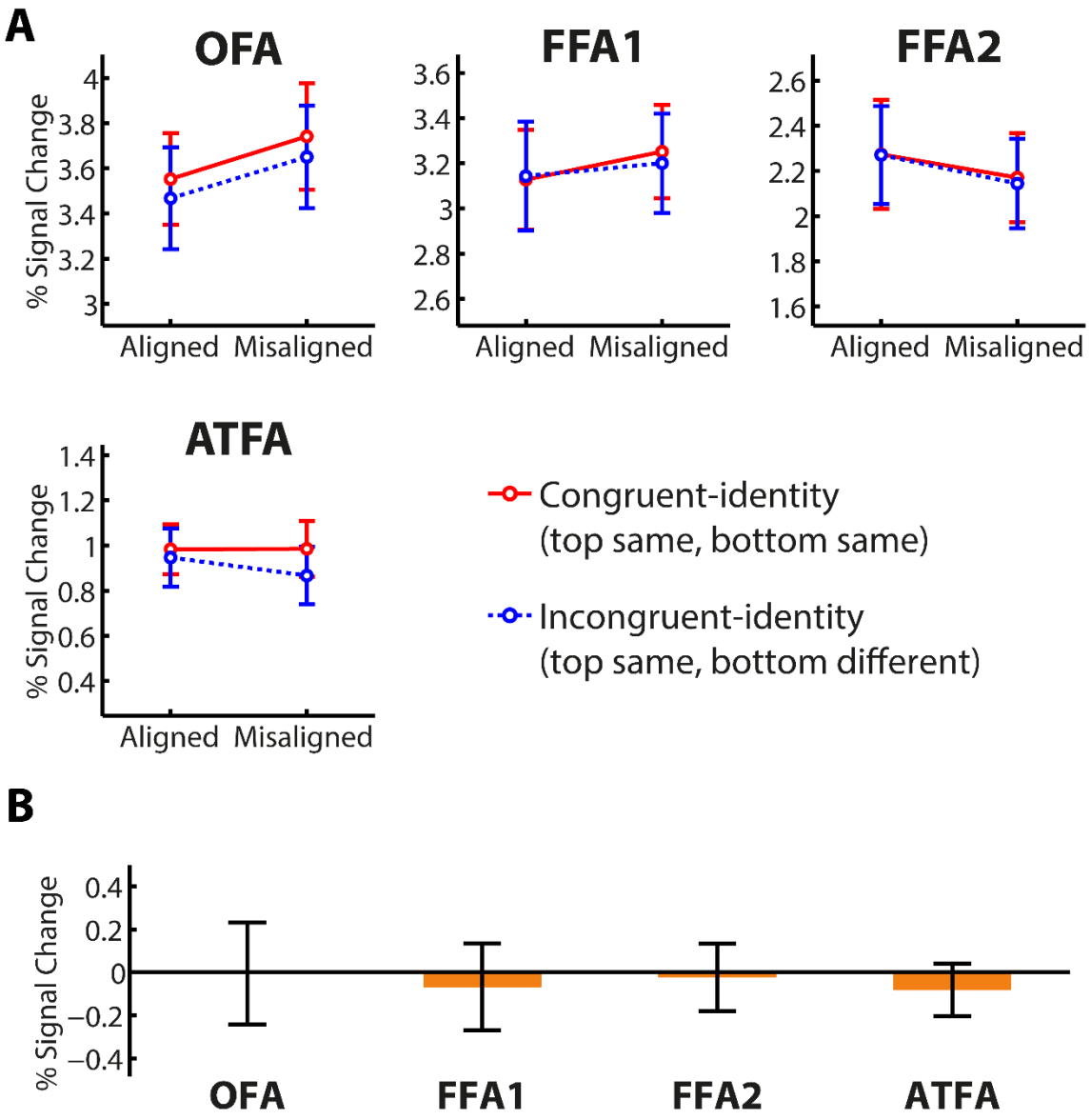

Figure 8. Neural responses to the top-different conditions in the face-responsive ROIs. (A) % signal change as a function of congruency and alignment. (B) Interaction between congruency and alignment with the contrast: *(aligned congruent-identity – aligned incongruent-identity) – (misaligned congruent-identity – misaligned incongruent-identity)*. Error bars indicate ±1 *SEM*.

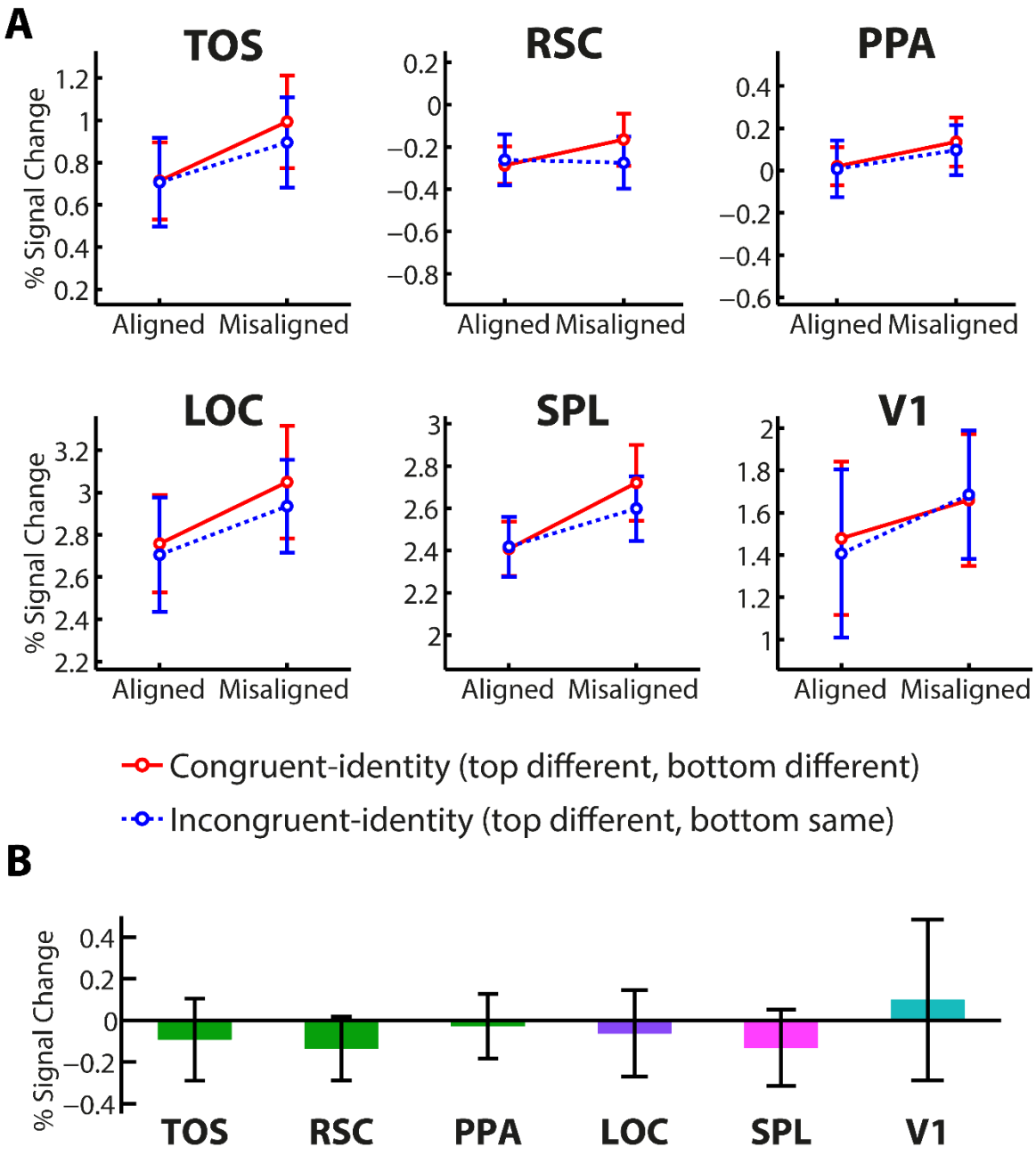

Figure 9. Neural responses to the top-different conditions in the scene-responsive, object-responsive, perceptual grouping and early visual ROIs. (A) % signal change as a function of congruency and alignment. (B) Interaction between congruency and alignment with the contrast: *(aligned congruent-identity – aligned incongruent-identity) – (misaligned congruent-identity – misaligned incongruent-identity)*. Error bars indicate ±1 *SEM*.

5.3.3. Correlation between behaviour and neural activity

We tested whether there was a correlation between the strength of participants' neural activity and behaviour related to holistic processing. In some cases, the strength of reaction time difference scores are known to have poor retest reliability for the same subject tested multiple times (Draheim et al., 2019). Therefore, we tested whether participants showed a correlation between their interaction effect measured with reaction times across the three experimental runs, to ensure that participants showed a consistent strength for this effect. We found a strong correlation between participants interaction effect measured with reaction times across runs (Run 1 and 2: r = 0.83, p = 1.3 x 10^{-5}; Run 2 and 3: r = 0.87, p = 9.7 x 10^{-7}; Run 1 and 3: r = 0.59, p = 0.0083), demonstrating that participants showed a consistent strength in their interaction effect measured with reaction times across the experiment.

5.3.3.1. Correlation with neural responses to alignment

We first tested whether the strength of the difference in neural responses to aligned versus misaligned congruent faces that we identified in the FFA2, LOC and TOS (Fig. 5) was associated with the strength of the interaction effect between alignment and congruency as measured behaviourally with accuracy and reaction times (Fig. 3). As illustrated in Fig. 10, we found no significant brain-behaviour correlation between the alignment effect in any ROI and the behavioural interaction effect (for the top-same conditions) with either response accuracy (FFA2: r = 0.12, p = 0.65; LOC: r = 0.016, p = 0.95; TOS: r = 0.088, p = 0.72) or reaction times (FFA2: r = 0.13, p = 0.62; LOC: r = 0.24, p = 0.33; TOS: r = 0.22, p = 0.36).

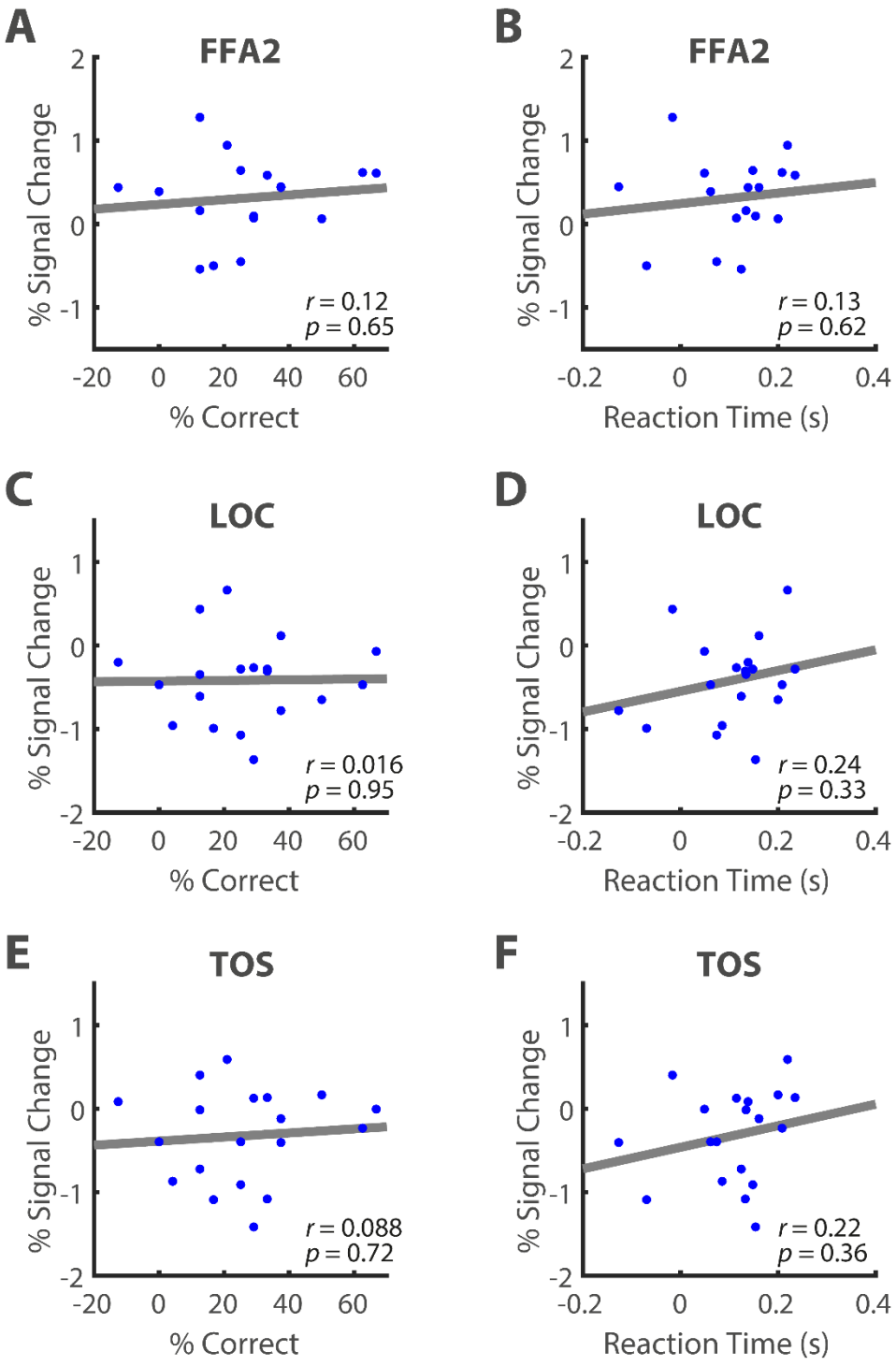

Figure 10. Correlation of differences in neural responses to alignment with behavioural measures of holistic processing in the top-same conditions. (A), (C) and (E) show the correlation between the interaction effect of congruency and alignment measured with response accuracy (% correct) and the difference in neural responses to aligned versus misaligned faces in the FFA2 (A), LOC (C) and TOS (E). (B), (D) and (F) show the correlation between the interaction effect of congruency and

alignment measured with reaction times and the difference in neural responses to aligned versus misaligned faces in the FFA2 (B), LOC (D) and TOS (F).

5.3.3.2. Correlation with congruency and alignment interaction

Secondly, we tested whether the strength of the interaction effect between alignment and congruency that we identified in the RSC and PPA (Fig. 8) was associated with the strength of the interaction effect between alignment and congruency as measured behaviourally with accuracy and reaction times (Fig. 3). As illustrated in Fig. 11, for the interaction effect measured with response accuracy, there was no significant brain-behaviour correlation in either RSC (r = -0.15, p = 0.56) or PPA (r = 0.20, p = 0.40). In contrast, for the interaction effect measured with reaction times, we found significant correlations between the behavioural holistic processing effect and the neural holistic processing effect observed in both RSC (r = 0.61, p = 0.0074) and PPA (r = 0.61, p = 0.0056), both surviving Bonferroni-correction for N = 2 ROIs.

We note that three of our participants showed a low or negative trend in both their neural interaction effect in the RSC and PPA and their behavioural interaction effect measured with reaction times. Therefore, we investigated if these participants were driving the brain-behaviour correlation by investigating the brain-behaviour correlation effect with these participants removed. We found a trend in the PPA (r = 0.43, p = 0.098), but no effect in the RSC (r = 0.12, p = 0.66). This suggests that our correlation result could be driven to some extent by two subgroups of participants, one subgroup that shows an interaction effect between alignment and congruency in both their neural responses (in the RSC and PPA) and in their reaction times, and a second subgroup that does not show either effect.

We additionally performed whole brain analyses to investigate if any other brain regions showed a correlation between neural activity and behaviour related to holistic processing. No regions were identified in these analyses.

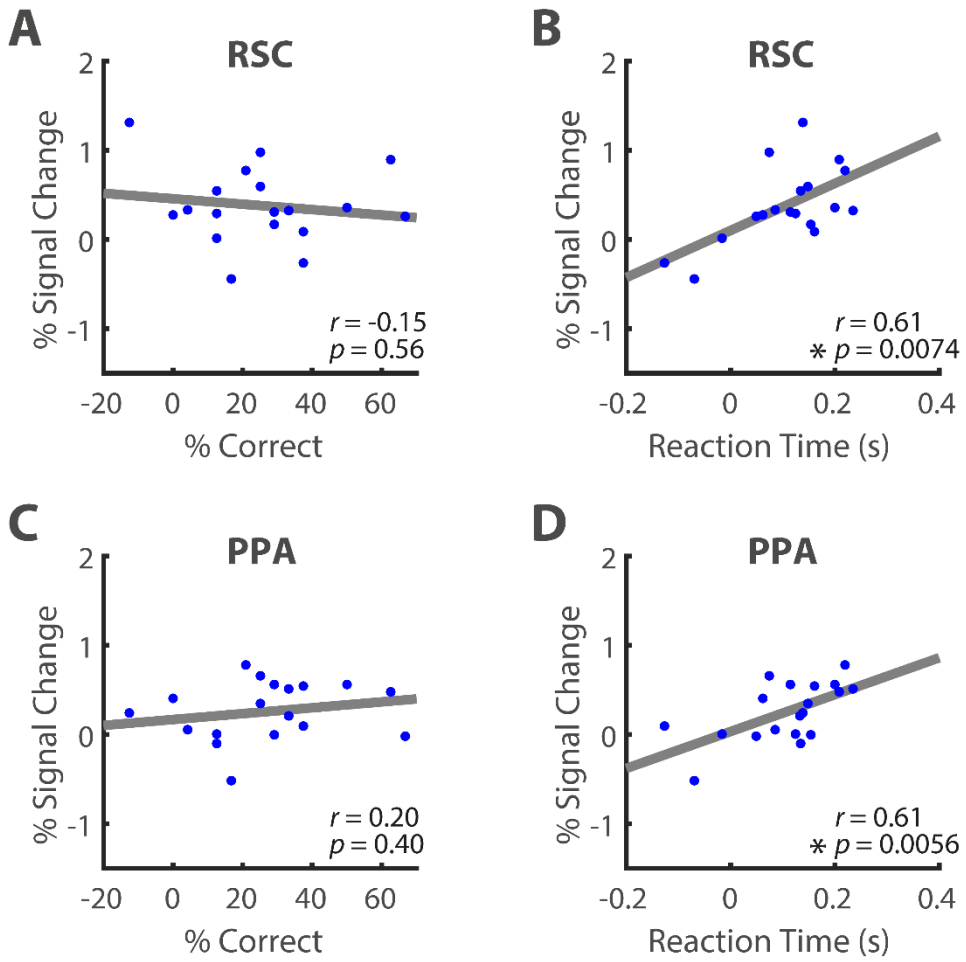

Figure 11. Correlation of interaction effect of congruency and alignment in the top-same conditions measured with neural responses and behavioural measures. (A) and (C) show the correlation between the interaction effect of congruency and alignment measured with response accuracy (% correct) and neural responses in the RSC (A) and PPA (C). (B) and (D) show the correlation between the interaction effect of congruency and alignment measured with reaction times and neural responses in the RSC (B) and PPA (D). * indicates $p < 0.05$, Bonferroni-corrected for N = 2 ROIs.

5.4. Discussion

In this study, we investigated which brain regions show differences in neural responses related to holistic processing in the composite-face paradigm. We found differences between the neural responses to aligned and misaligned faces in the LOC, TOS and FFA2, suggesting that these regions are sensitive to how holistically faces are processed. Furthermore, we found that the RSC and PPA showed patterns of neural activation consistent with processing of face identity in a holistic manner. The strength of this effect in RSC and PPA directly correlated with participants' reaction times used as a behavioural measure of holistic processing. In combination, our results indicate that brain regions outside the commonly defined face-responsive network are involved in some aspects of holistic processing measured by the composite-face paradigm.

5.4.1. Neural responses to face alignment

In a first set of analyses, we investigated whether any brain regions show differences in neural activation between aligned and misaligned faces. We found significant differences in neural activation in the LOC, and marginal trends (that did not survive Bonferroni correction for N = 10 ROIs) in the TOS and FFA2. Both the LOC and TOS showed higher activation to misaligned faces compared to aligned faces, whereas the FFA2 showed higher activation to aligned faces compared to misaligned faces. Interestingly, our pattern of results shows similarity to the differences in neural activation between upright and inverted faces. Several studies have found higher activation to inverted compared to upright faces in the LOC (Aguirre et al., 1999; Epstein et al., 2006; Grotheer et al., 2014; Haxby et al., 1999). It has been proposed that this could be due to inverted faces being processed similarly to objects or due to a recruitment of these regions for more demanding face processing (Aguirre et al., 1999; Haxby et al., 1999). Our finding that responses in the LOC are higher to misaligned faces compared to aligned faces further demonstrates a sensitivity of the LOC to factors affecting holistic processing of faces.

Higher responses to aligned faces compared to misaligned faces in the FFA2 suggests that this region may be involved in processing faces holistically. This result is consistent with previous work that has shown that the FFA2 is involved in holistic processing of expertise objects (Ross et al., 2018), and with many previous studies that have linked activity in the

FFA to holistic processing of faces (Andrews et al., 2010; Goffaux et al., 2013; Schiltz et al., 2010; Schiltz and Rossion, 2006). Studies investigating neural responses to face inversion in the FFA have found mixed results. Some studies found higher responses to upright compared to inverted faces in the FFA (Goffaux et al., 2013; Yovel and Kanwisher, 2005), whereas some studies have did not find differences in responses (Aguirre et al., 1999; Epstein et al., 2006; Grotheer et al., 2014; Haxby et al., 1999). Based on our findings, we hypothesize that the different results of these studies could be due to differences in FFA localization, as we find a difference in neural responses in the FFA2 but not in the FFA1. The similarity between our results in the LOC and FFA2 and the findings for inverted faces suggests that misaligned faces may be processed similarly to inverted faces, both of which have been demonstrated to be processed less holistically than upright aligned faces in behaviour (Young et al., 1987).

5.4.2. Neural responses to composite-faces

In a second set of analyses, we investigated which brain regions show a difference in neural responses to congruent-identity and incongruent-identity trials of the composite-face paradigm that is reduced by misalignment. We predicted that any brain regions encoding face identity in a holistic manner would show this pattern of responses, due to an interference of the irrelevant bottom-half face information during the aligned conditions, but not misaligned conditions. We found neural responses consistent with this pattern in the RSC and PPA, two regions that have been shown to be important in spatial navigation, scene-processing, memory (Epstein, 2008; Epstein and Kanwisher, 1998; Vann et al., 2009) and other functions such as contextual relationships (Bar, 2004; Bar and Aminoff, 2003). We hypothesize that the sensitivity of these regions to the change in face identity induced by holistic processing in the composite-face paradigm could be related to the failure of selective spatial attention to the top-half of the face (due to holistic processing) that leads to this perceptual change. Furthermore, we hypothesize that that these regions could be recruited in spatially demanding face-tasks, such as the composite-face paradigm.

Surprisingly, we did not identify any differences in neural responses in the FFA related to the interaction between congruency and alignment, in contrast to several previous findings (Goffaux et al., 2013; Schiltz et al., 2010; Schiltz and Rossion, 2006). This difference in results may be due to differences in experimental factors (e.g. task, stimuli).

For example, it is known that repetition effects can sometimes differ across the duration of an experiment (Müller et al., 2013), thus it is possible that differences in the number of faces used in the stimulus set could lead to differences in ongoing repetition suppression effects. Our use of the complete-design version of the composite-face paradigm could be another reason we observe different results to the previous studies. We found higher BOLD responses when two faces were aligned and congruent (i.e. identical faces) compared to when they were aligned and incongruent (i.e. same top-halves, different bottom-halves) and this repetition-effect disappeared when the bottom halves of the faces were misaligned. Most studies find a repetition-suppression effect when subjects view two identical faces compared to when they view two different faces (Grill-Spector et al., 1999), although repetition-enhancement was also found in many studies (Segaert et al., 2013). We believe the reason we find a repetition-enhancement effect in this study is due to our use of the complete-design version of the composite-face paradigm, and the role of expectation in fMRI repetition effects. High-level top down influences have been shown to modulate how stimulus repetition affects evoked neural activity. For example, the probability of repetitions occurring in an experimental run changes the repetition-effect strength (Larsson and Smith, 2012; Summerfield et al., 2008). In both FFA and PPA the response to faces was reduced when subjects had higher expectation of seeing a face, compared to lower expectation of seeing a face (Egner et al., 2010) In our experiment, the use of the complete design meant that there were more trials where subjects perceived the top-halves of the faces to be different compared to trials where they perceived them to be the same (our experimental design had equal numbers of top-same and top-different conditions, but the composite-illusion leads to incongruent-identity aligned top-same stimuli being perceived as different). Thus, subjects may expect top-different trials more than top-same ones, leading them to have a lower BOLD response when they perceive the faces to be different compared to when they perceive them to be the same. Importantly, the repetition-effect pattern was in accordance with our expected pattern related to holistic processing (i.e. difference between aligned congruent-identity and incongruent-identity conditions, and no difference between misaligned congruent-identity and incongruent-identity conditions).

5.4.3. Linking holistic processing in neural activity and behaviour

We found that the strength of the interaction effect between congruency and alignment in the RSC and PPA correlated with the behavioural interaction effect between congruency and alignment measured with reaction times, but not with accuracy. Could this correlation with reaction times reflect differences in time on task or attentional differences driving the differences in neural responses in RSC and PPA? We hypothesize that this is unlikely as greater time on task or stronger attention would both be likely to lead to higher neural activation rather than the lower neural activations we find in RSC and PPA for incongruent-identity conditions compared to congruent-identity conditions. Why might we find a correlation with reaction times but not with accuracy? It has been suggested that reaction times may be particularly important as a composite-effect behavioural measure when stimulus presentation times are longer, as in the present study (Rossion, 2013). The correlation between holistic processing as measured with reaction times and the holistic processing effect measured in RSC and PPA suggests there is a direct link between the brain activity we recorded in these areas and the behavioural performance of the participants in the composite-task. In a whole-brain analysis, no other brain area was related to holistic processing and behavioural measures. Furthermore, we did not identify any significant correlation between participants' behavioural responses and the differences in their neural responses to aligned and misaligned faces.

We separately investigated the neural activity and behavioural responses to the top-same and top-different conditions in this study. The necessity of incongruent-identity top-different conditions for measuring holistic face processing with the composite paradigm is debated (Richler and Gauthier, 2013; Rossion, 2013). Furthermore, previous work has suggested that evidence of holistic face processing is mainly found when the target face part being matched is more similar rather than different (Goffaux, 2012; Goffaux et al., 2013). In our study, we found differences in neural responses during the top-same but not the top-different conditions of the composite face paradigm, consistent with these previous findings.

5.4.4. Holistic processing outside of face-responsive regions

In both of our fMRI analyses we found neural responses related to holistic processing in brain regions outside of the face-responsive network. Previous behavioural studies have shown that objects can be processed holistically (Bukach et al., 2010; Diamond and Carey, 1986; Gauthier and Tarr, 1997; Wong et al., 2009a), and that there is interference between holistic processing of objects and faces in behaviour (Curby and Gauthier, 2014; Curby and Moerel, 2019; Gauthier et al., 2003) suggesting that there is some overlap in the neural mechanisms involved in holistic processing of faces and objects. Our results suggest there could be some overlap in neural processing outside of the face-responsive network. Furthermore, our results may shed light on why some prosopagnosic individuals have intact holistic processing ability (Biotti et al., 2017; Le Grand et al., 2006; Susilo et al., 2010), why face recognition ability and strength of holistic processing are not necessarily correlated (Konar et al., 2010; Zhao et al., 2014b), but see also (Richler et al., 2011a), and why factors affecting face recognition ability (e.g. contrast negation) can leave holistic processing unaffected (Hole et al., 1999; Taubert and Alais, 2011). Previous studies have shown that face-responsive ATFA and FFA are involved in face recognition (Grill-Spector et al., 2004; Nasr and Tootell, 2012), thus if holistic face processing occurs in more disperse high-level occipital-temporal regions, it is not surprising that face recognition and holistic face processing may in some cases not be linked, as the neural processes involved would not completely overlap.

5.4.5. Conclusion

We investigated neural responses related to holistic face processing, using the composite paradigm in functionally defined brain regions related to face, scene, object and Gestalt processing. We identified differences in neural responses to aligned and misaligned faces in the LOC, TOS and FFA2, suggesting that these regions are sensitive to holistic processing of faces. Furthermore, we found neural activity consistent with processing face identity in a holistic manner in the RSC and PPA, and found that the strength of this activity correlated with participants' behaviour in the composite face paradigm. Our results show the importance of investigating face processing mechanisms in a wide range of brain regions, and suggest that regions both within and outside of the face-responsive brain network are involved in holistic face processing.

Acknowledgements

This research was supported by the Max Planck Society, Germany.

Declarations of interest

None.

References

Aguirre, G.K., 2007. Continuous carry-over designs for fMRI. Neuroimage 35, 1480–1494. https://doi.org/10.1016/j.neuroimage.2007.02.005

Aguirre, G.K., Singh, R., D'Esposito, M., 1999. Stimulus inversion and the responses of face and object-sensitive cortical areas. Neuroreport 10, 189–194. https://doi.org/10.1097/00001756-199901180-00036

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. J. Neurosci. 30, 3544–3552. https://doi.org/10.1523/JNEUROSCI.4863-09.2010

Arcurio, L.R., Gold, J.M., James, T.W., 2012. The response of face-selective cortex with single face parts and part combinations. Neuropsychologia 50, 2454–2459. https://doi.org/10.1016/j.neuropsychologia.2012.06.016

Bar, M., 2004. Visual objects in context. Nat. Rev. Neurosci. 5, 617–629. https://doi.org/10.1038/nrn1476

Bar, M., Aminoff, E., 2003. Cortical analysis of visual context. Neuron 38, 347–358. https://doi.org/10.1016/S0896-6273(03)00167-3

Bettencourt, K.C., Xu, Y., 2013. The Role of Transverse Occipital Sulcus in Scene Perception and Its Relationship to Object Individuation in Inferior Intraparietal Sulcus. J. Cogn. Neurosci. 25, 1711–1722. https://doi.org/10.1162/jocn_a_00422

Biotti, F., Wu, E., Yang, H., Jiahui, G., Duchaine, B., Cook, R., 2017. Normal composite face effects in developmental prosopagnosia. Cortex 95, 63–76. https://doi.org/10.1016/j.cortex.2017.07.018

Blanz, V., Vetter, T., 1999. A morphable model for the synthesis of 3D faces. SIGGRAPH'99 Conf. Proc. 187–194. https://doi.org/10.1145/311535.311556

Brainard, D.H., 1997. The Psychophysics Toolbox. Spat. Vis. 10, 433–436. https://doi.org/10.1163/156856897X00357

Brandman, T., Yovel, G., 2016. Bodies are Represented as Wholes Rather Than Their Sum of Parts in the Occipital-Temporal Cortex. Cereb. Cortex 26, 530–543. https://doi.org/10.1093/cercor/bhu205

Brooks, J.L., 2012. Counterbalancing for serial order carryover effects in experimental condition orders. Psychol. Methods 17, 600–614. https://doi.org/10.1037/a0029310

Bukach, C.M., Phillips, W.S., Gauthier, I., 2010. Limits of generalization between categories and implications for theories of category specificity. Attention, Perception, Psychophys. 72, 1865–1874. https://doi.org/10.3758/APP.72.7.1865

Chua, K.-W., Gauthier, I., 2020. Domain-Specific Experience Determines Individual Differences in Holistic Processing. J. Exp. Psychol. Gen. 149, 31–41. https://doi.org/10.1037/xge0000628

Curby, K.M., Gauthier, I., 2014. Interference between face and non-face domains of perceptual expertise: a replication and extension. Front. Psychol. 5, 1–11. https://doi.org/10.3389/fpsyg.2014.00955

Curby, K.M., Huang, M., Moerel, D., 2019. Multiple paths to holistic processing: Holistic processing of Gestalt stimuli do not overlap with holistic face processing in the same manner as do objects of expertise. Attention, Perception, Psychophys. 81, 716–726. https://doi.org/10.3758/s13414-018-01643-x

Curby, K.M., Moerel, D., 2019. Behind the face of holistic perception: Holistic processing of Gestalt stimuli and faces recruit overlapping perceptual mechanisms. Attention, Perception, Psychophys. 81, 2873–2880. https://doi.org/10.3758/s13414-019-01749-w

Diamond, R., Carey, S., 1986. Why faces are and are not special: an effect of expertise. J. Exp. Psychol. Gen. 115, 107–117. https://doi.org/10.1037//0096-3445.115.2.107

Dilks, D.D., Julian, J.B., Paunov, A.M., Kanwisher, N., 2013. The Occipital Place Area Is Causally and Selectively

Involved in Scene Perception. J. Neurosci. 33, 1331–1336. https://doi.org/10.1523/JNEUROSCI.4081-12.2013

Draheim, C., Mashburn, C.A., Martin, J.D., Engle, R.W., 2019. Reaction time in differential and developmental research: A review and commentary on the problems and alternatives. Psychol. Bull. 145, 508–535. https://doi.org/10.1037/bul0000192

Egner, T., Monti, J.M., Summerfield, C., 2010. Expectation and Surprise Determine Neural Population Responses in the Ventral Visual Stream. J. Neurosci. 30, 16601–16608. https://doi.org/10.1523/JNEUROSCI.2770-10.2010

Epstein, R., 2008. Parahippocampal and retrosplenial contributions to human spatial navigation. Trends Cogn. Sci. 12, 388–396. https://doi.org/10.1016/j.tics.2008.07.004

Epstein, R., Higgins, J.S., Parker, W., Aguirre, G.K., Cooperman, S., 2006. Cortical correlates of face and scene inversion: A comparison. Neuropsychologia 44, 1145–1158. https://doi.org/10.1016/j.neuropsychologia.2005.10.009

Epstein, R., Kanwisher, N., 1998. A cortical representation of the local visual environment. Nature 392, 598–601. https://doi.org/10.1038/33402

Farah, M.J., Wilson, K.D., Drain, M., Tanaka, J.N., 1998. What is "special" about face perception? Psychol. Rev. 105, 482–498. https://doi.org/10.1037/0033-295X.105.3.482

Faul, F., Erdfelder, E., Lang, A.G., Buchner, A., 2007. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav. Res. Methods 39, 175–191. https://doi.org/10.3758/BF03193146

Friston, K.J., Rotshtein, P., Geng, J.J., Sterzer, P., Henson, R.N., 2006. A critique of functional localisers. Neuroimage 30, 1077–1087. https://doi.org/10.1016/j.neuroimage.2005.08.012

Gauthier, I., Curran, T., Curby, K.M., Collins, D., 2003. Perceptual interference supports a non-modular account of face processing. Nat. Neurosci. 6, 428–432. https://doi.org/10.1038/nn1029

Gauthier, I., Skudlarski, P., Gore, J.C., Anderson, A.W., 2000a. Expertise for cars and birds recruits brain areas involved in face recognition. Nat. Neurosci. 3, 191–197. https://doi.org/10.1038/72140

Gauthier, I., Tarr, M.J., 2002. Unraveling mechanisms for expert object recognition: bridging brain activity and behavior. J. Exp. Psychol. Hum. Percept. Perform. 28, 431–446. https://doi.org/10.1037/0096-1523.28.2.431

Gauthier, I., Tarr, M.J., 1997. Becoming a "Greeble" expert: Exploring mechanisms for face recognition. Vision Res. 37, 1673–1682. https://doi.org/10.1016/S0042-6989(96)00286-6

Gauthier, I., Tarr, M.J., Moylan, J., Skudlarski, P., Gore, J.C., Anderson, A.W., 2000b. The fusiform "face area" is part of a network that processes faces at the individual level. J. Cogn. Neurosci. 12, 495–504. https://doi.org/10.1162/089892900562165

Goffaux, V., 2012. The discriminability of local cues determines the strength of holistic face processing. Vision Res. 64, 17–22. https://doi.org/10.1016/j.visres.2012.04.022

Goffaux, V., Schiltz, C., Mur, M., Goebel, R., 2013. Local discriminability determines the strength of holistic processing for faces in the fusiform face area. Front. Psychol. 3, 1–14. https://doi.org/10.3389/fpsyg.2012.00604

Grassi, P.R., Zaretskaya, N., Bartels, A., 2018. A Generic Mechanism for Perceptual Organization in the Parietal Cortex. J. Neurosci. 38, 7158–7169. https://doi.org/10.1523/JNEUROSCI.0436-18.2018

Grassi, P.R., Zaretskaya, N., Bartels, A., 2016. Parietal cortex mediates perceptual Gestalt grouping independent of stimulus size. Neuroimage 133, 367–377. https://doi.org/10.1016/j.neuroimage.2016.03.008

Grill-Spector, K., 2003. The neural basis of object perception. Curr. Opin. Neurobiol. 13, 159–166.

https://doi.org/10.1016/S0959-4388(03)00040-0

Grill-Spector, K., Knouf, N., Kanwisher, N., 2004. The fusiform face area subserves face perception, not generic within-category identification. Nat. Neurosci. 7, 555–562. https://doi.org/10.1038/nn1224

Grill-Spector, K., Kushnir, T., Edelman, S., Avidan, G., Itzchak, Y., Malach, R., 1999. Differential processing of objects under various viewing conditions in the human lateral occipital complex. Neuron 24, 187–203. https://doi.org/10.1016/S0896-6273(00)80832-6

Grotheer, M., Hermann, P., Vidnyánszky, Z., Kovács, G., 2014. Repetition probability effects for inverted faces. Neuroimage 102, 416–423. https://doi.org/10.1016/j.neuroimage.2014.08.006

Harris, A., Aguirre, G.K., 2010. Neural Tuning for Face Wholes and Parts in Human Fusiform Gyrus Revealed by fMRI Adaptation. J. Neurophysiol. 104, 336–345. https://doi.org/10.1152/jn.00626.2009

Harris, A., Aguirre, G.K., 2008. The representation of parts and wholes in face-selective cortex. J. Cogn. Neurosci. 20, 863–78. https://doi.org/10.1162/jocn.2008.20509

Haxby, J. V., Ungerleider, L.G., Clark, V.P., Schouten, J.L., Hoffman, E.A., Martin, A., 1999. The effect of face inversion on activity in human neural systems for face and object perception. Neuron 22, 189–199. https://doi.org/10.1016/S0896-6273(00)80690-X

Hinds, O., Polimeni, J.R., Rajendran, N., Balasubramanian, M., Amunts, K., Zilles, K., Schwartz, E.L., Fischl, B., Triantafyllou, C., 2009. Locating the functional and anatomical boundaries of human primary visual cortex. Neuroimage 46, 915–922. https://doi.org/10.1016/j.neuroimage.2009.03.036

Hole, G.J., 1994. Configurational factors in the perception of unfamiliar faces. Perception 23, 65–74. https://doi.org/10.1068/p230065

Hole, G.J., George, P.A., Dunsmore, V., 1999. Evidence for holistic processing of faces viewed as photographic negatives. Perception 28, 341–359. https://doi.org/10.1068/p2622

Kamps, F.S., Julian, J.B., Kubilius, J., Kanwisher, N., Dilks, D.D., 2016. The occipital place area represents the local elements of scenes. Neuroimage 132, 417–424. https://doi.org/10.1016/j.neuroimage.2016.02.062

Kanwisher, N., McDermott, J., Chun, M.M., 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. J. Neurosci. 17, 4302–4311.

Kleiner, M., Brainard, D., Pelli, D., 2007. "What's new in Psychtoolbox-3?," in: Perception 36 ECVP Abstract Supplement.

Konar, Y., Bennett, P.J., Sekuler, A.B., 2010. Holistic processing is not correlated with face-identification accuracy. Psychol. Sci. 21, 38–43. https://doi.org/10.1177/0956797609356508

Larsson, J., Smith, A.T., 2012. fMRI repetition suppression: Neuronal adaptation or stimulus expectation? Cereb. Cortex 22, 567–576. https://doi.org/10.1093/cercor/bhr119

Le Grand, R., Cooper, P.A., Mondloch, C.J., Lewis, T.L., Sagiv, N., de Gelder, B., Maurer, D., 2006. What aspects of face processing are impaired in developmental prosopagnosia? Brain Cogn. 61, 139–158. https://doi.org/10.1016/j.bandc.2005.11.005

Liu, J., Harris, A., Kanwisher, N., 2010. Perception of face parts and face configurations: an fMRI study. J. Cogn. Neurosci. 22, 203–211. https://doi.org/10.1162/jocn.2009.21203

Maguire, E., 2001. The retrosplenial contribution to human navigation: A review of lesion and neuroimaging findings. Scand. J. Psychol. 42, 225–238. https://doi.org/10.1111/1467-9450.00233

Malach, R., Reppas, J.B., Benson, R.R., Kwong, K.K., Jiang, H., Kennedy, W.A., Ledden, P.J., Brady, T.J., Rosen, B.R., Tootell, R.B.H., 1995. Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. Proc. Natl. Acad. Sci. U. S. A. 92, 8135–8139. https://doi.org/10.1073/pnas.92.18.8135

Maurer, D., Le Grand, R., Mondloch, C.J., 2002. The many faces of configural processing. Trends Cogn. Sci. 6,

255–260.

Müller, N.G., Strumpf, H., Scholz, M., Baier, B., Melloni, L., 2013. Repetition suppression versus enhancement - It's quantity that matters. Cereb. Cortex 23, 315–322. https://doi.org/10.1093/cercor/bhs009

Nasr, S., Tootell, R.B.H., 2012. Role of fusiform and anterior temporal cortical areas in facial recognition. Neuroimage 63, 1743–1753. https://doi.org/10.1016/j.neuroimage.2012.08.031

Rajimehr, R., Young, J.C., Tootell, R.B.H., 2009. An anterior temporal face patch in human cortex, predicted by macaque maps. Proc. Natl. Acad. Sci. U. S. A. 106, 1995–2000. https://doi.org/10.1073/pnas.0807304106

Richler, J.J., Cheung, O.S., Gauthier, I., 2011a. Holistic processing predicts face recognition. Psychol. Sci. 22, 464–471. https://doi.org/10.1177/0956797611401753

Richler, J.J., Gauthier, I., 2014. A meta-analysis and review of holistic face processing. Psychol. Bull. 140, 1281–1302. https://doi.org/10.1037/a0037004

Richler, J.J., Gauthier, I., 2013. When intuition fails to align with data: A reply to Rossion (2013). Vis. cogn. 21, 254–276. https://doi.org/10.1080/13506285.2013.796035

Richler, J.J., Mack, M.L., Palmeri, T.J., Gauthier, I., 2011b. Inverted faces are (eventually) processed holistically. Vision Res. 51, 333–342. https://doi.org/10.1016/j.visres.2010.11.014

Ross, D.A., Tamber-Rosenau, B.J., Palermi, T.J., Zhang, J., Xu, Y., Gauthier, I., 2018. High-resolution Functional Magnetic Resonance Imaging Reveals Configural Processing of Cars in Right Anterior Fusiform Face Area of Car Experts. J. Cogn. Neurosci. 30, 973–984. https://doi.org/10.1162/jocn_a_01256

Rossion, B., 2013. The composite face illusion: A whole window into our understanding of holistic face perception. Vis. cogn. 21, 139–253. https://doi.org/10.1080/13506285.2013.772929

Rossion, B., Boremanse, A., 2008. Nonlinear relationship between holistic processing of individual faces and picture-plane rotation: Evidence from the face composite illusion. J. Vis. 8, 1–13. https://doi.org/10.1167/8.4.3

Saxe, R., Brett, M., Kanwisher, N., 2006. Divide and conquer: A defense of functional localizers. Neuroimage 30, 1088–1096. https://doi.org/10.1016/j.neuroimage.2005.12.062

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. J. Vis. 10, 1–16. https://doi.org/10.1167/10.2.25

Schiltz, C., Rossion, B., 2006. Faces are represented holistically in the human occipito-temporal cortex. Neuroimage 32, 1385–1394. https://doi.org/10.1016/j.neuroimage.2006.05.037

Schindler, A., Bartels, A., 2016. Visual high-level regions respond to high-level stimulus content in the absence of low-level confounds. Neuroimage 132, 520–525. https://doi.org/10.1016/j.neuroimage.2016.03.011

Segaert, K., Weber, K., de Lange, F.P., Petersson, K.M., Hagoort, P., 2013. The suppression of repetition enhancement: A review of fMRI studies. Neuropsychologia 51, 59–66. https://doi.org/10.1016/j.neuropsychologia.2012.11.006

Summerfield, C., Trittschuh, E.H., Monti, J.M., Mesulam, M.M., Egner, T., 2008. Neural repetition suppression reflects fulfilled perceptual expectations. Nat. Neurosci. 11, 1004–1006. https://doi.org/10.1038/nn.2163

Susilo, T., McKone, E., Dennett, H., Darke, H., Palermo, R., Hall, A., Pidcock, M., Dawel, A., Jeffery, L., Wilson, C.E., Rhodes, G., 2010. Face recognition impairments despite normal holistic processing and face space coding: Evidence from a case of developmental prosopagnosia. Cogn. Neuropsychol. 27, 636–664. https://doi.org/10.1080/02643294.2011.613372

Tanaka, J.W., Farah, M.J., 1993. Parts and Wholes in Face Recognition. Q. J. Exp. Psychol. Sect. A 46, 225–245. https://doi.org/10.1080/14640749308401045

Taubert, J., Alais, D., 2011. Identity aftereffects, but not composite effects, are contingent on contrast polarity. Perception 40, 422–436. https://doi.org/10.1068/p6874

Troje, N.F., Bülthoff, H.H., 1996. Face recognition under varying poses: The role of texture and shape. Vision Res. 36, 1761–1771. https://doi.org/10.1016/0042-6989(95)00230-8

Tsao, D.Y., Moeller, S., Freiwald, W.A., 2008. Comparing face patch systems in macaques and humans. Proc. Natl. Acad. Sci. U. S. A. 105, 19514–19519. https://doi.org/10.1073/pnas.0809662105

Vann, S.D., Aggleton, J.P., Maguire, E.A., 2009. What does the retrosplenial cortex do? Nat. Rev. Neurosci. 10, 792–802. https://doi.org/10.1038/nrn2733

Weiner, K.S., Golarai, G., Caspers, J., Chuapoco, M.R., Mohlberg, H., Zilles, K., Amunts, K., Grill-Spector, K., 2014. The mid-fusiform sulcus: A landmark identifying both cytoarchitectonic and functional divisions of human ventral temporal cortex. Neuroimage 84, 453–465. https://doi.org/10.1016/j.neuroimage.2013.08.068

Weiner, K.S., Grill-Spector, K., 2013. Neural representations of faces and limbs neighbor in human high-level visual cortex: Evidence for a new organization principle. Psychol. Res. 77, 74–97. https://doi.org/10.1007/s00426-011-0392-x

Weiner, K.S., Jonas, J., Gomez, J., Maillard, L., Brissart, H., Hossu, G., Jacques, C., Loftus, D., Colnat-Coulbois, S., Stigliani, A., Barnett, M.A., Grill-Spector, K., Rossion, B., 2016. The Face-Processing Network Is Resilient to Focal Resection of Human Visual Cortex. J. Neurosci. 36, 8425–8440. https://doi.org/10.1523/JNEUROSCI.4509-15.2016

Wong, A.C.-N., Palmeri, T.J., Gauthier, I., 2009a. Conditions for facelike expertise with objects: Becoming a ziggerin expert - but which type? Psychol. Sci. 20, 1108–1117. https://doi.org/10.1111/j.1467-9280.2009.02430.x

Wong, A.C.-N., Palmeri, T.J., Rogers, B.P., Gore, J.C., Gauthier, I., 2009b. Beyond shape: How you learn about objects affects how they are represented in visual cortex. PLoS One 4, 1–7. https://doi.org/10.1371/journal.pone.0008405

Xu, Y., 2005. Revisiting the role of the fusiform face area in visual expertise. Cereb. Cortex 15, 1234–1242. https://doi.org/10.1093/cercor/bhi006

Young, A.W., Hellawell, D., Hay, D.C., 1987. Configurational information in face perception. Perception 16, 747–759. https://doi.org/10.1068/p160747n

Yovel, G., Kanwisher, N., 2005. The neural basis of the behavioral face-inversion effect. Curr. Biol. 15, 2256–2262. https://doi.org/10.1016/j.cub.2005.10.072

Zachariou, V., Nikas, C. V., Safiullah, Z.N., Gotts, S.J., Ungerleider, L.G., 2017. Spatial mechanisms within the dorsal visual pathway contribute to the configural processing of faces. Cereb. Cortex 27, 4124–4138. https://doi.org/10.1093/cercor/bhw224

Zaretskaya, N., Anstis, S., Bartels, A., 2013. Parietal Cortex Mediates Conscious Perception of Illusory Gestalt. J. Neurosci. 33, 523–531. https://doi.org/10.1523/JNEUROSCI.2905-12.2013

Zhao, M., Bülthoff, H.H., Bülthoff, I., 2016. Beyond faces and expertise: Face-like holistic processing of nonface objects in the absence of expertise. Psychol. Sci. 27, 213–222. https://doi.org/10.1177/0956797615617779

Zhao, M., Bülthoff, I., 2017. Holistic processing of static and moving faces. J. Exp. Psychol. Learn. Mem. Cogn. 43, 1020–1035. https://doi.org/10.1037/xlm0000368

Zhao, M., Cheung, S.-H., Wong, A.C.-N., Rhodes, G., Chan, E.K.S., Chan, W.W.L., Hayward, W.G., 2014a. Processing of configural and componential information in face-selective cortical areas. Cogn. Neurosci. 5, 160–167. https://doi.org/10.1080/17588928.2014.912207

Zhao, M., Hayward, W.G., Bülthoff, I., 2014b. Holistic processing, contact, and the other-race effect in face recognition. Vision Res. 105, 61–69. https://doi.org/10.1016/j.visres.2014.09.006

6. No holistic processing of objects in brain regions that process faces holistically, despite an identical behavioural effect

Abstract

Behavioural studies have demonstrated that faces and objects of expertise are processed holistically and have linked this holistic processing to brain activity in occipitotemporal cortical regions, in particular the fusiform face area. Recent behavioural work has shown that non-expertise objects with salient Gestalt information are also processed holistically. It remains unclear which brain areas are involved in holistic processing of these non-expertise objects. In this study, participants performed a composite task with non-expertise line pattern objects, while we recorded their brain activity with fMRI. These line pattern objects have previously been shown to elicit strong evidence of holistic processing in behaviour. We defined brain regions of interest based on their responses to objects, scenes, faces and perceptual grouping and investigated how activity in these regions related to holistic processing in a composite task. Despite our participants showing strong evidence of holistic processing in their behavioural responses during the fMRI experiment, we found that neither regions shown to process faces holistically (in previous studies and our own work in Chapter 5), nor any other brain regions we investigated, showed activity consistent with holistic processing. We conclude that different brain regions may underlie holistic processing of faces and non-expertise objects, but further work is needed to elucidate which brain regions underlie holistic processing of non-expertise objects.

Keywords: holistic processing, composite-face effect, object, expertise

6.1. Introduction

Holistic processing refers to the tendency to perceive an object as an indecomposable whole, rather than as a collection of independent parts. Faces are the class of objects that has been most consistently demonstrated to be processed holistically. One behavioural measure that has been used to demonstrate holistic processing of faces is the composite-face paradigm. In this paradigm, when the top-half of one person's face is aligned with the bottom-halves of different faces (i.e. composite faces), observers have the tendency to perceive the top-halves of the faces as different identities, as they are unable to ignore the bottom halves of the composite faces (Hole, 1994; Young, Hellawell, & Hay, 1987). Another behavioural measure that has been used to demonstrate evidence of holistic processing is the part-whole paradigm. Observers learn to recognise the faces of different individuals, and then are tested on their ability to recognise a part of a face (e.g. the nose) either alone or in the context of the whole face. Observers tend to show an advantage for recognising the face part in the whole face context (Tanaka & Farah, 1993). Interestingly, the evidence of holistic processing from both of these paradigms has been shown to be significantly reduced when faces are shown in an inverted compared to an upright configuration (Tanaka & Farah, 1993; Young et al., 1987), showing that holistic face processing is strongest for upright faces.

In addition to faces, behavioural studies have shown that expertise in recognising objects can also lead to these objects being processed holistically. For example, car experts tested with a composite paradigm show a similar inability to ignore the irrelevant bottom half of composite car images (Bukach, Phillips, & Gauthier, 2010) and dog experts, but not novices, show a deficit when recognising dogs from inverted images compared to upright ones (Diamond & Carey, 1986). Furthermore, participants that have been trained in the lab to become experts at recognising novel classes of objects (that must be recognised by the specific arrangement of their parts) have been shown to develop holistic processing of these objects (Gauthier & Tarr, 1997; Wong, Palmeri, & Gauthier, 2009). The authors proposed that holistic processing may occur with the development of expertise in individuating objects (Wong et al., 2009). However, more recent work has shown that other factors, such as Gestalt information may also be involved in holistic processing of objects. Non-expertise

line pattern objects, containing salient Gestalt information, were found to be processed as holistically as faces in a composite paradigm (Zhao, Bülthoff, & Bülthoff, 2016), and weakening the Gestalt information in these line pattern objects led to a decrease in the evidence of holistic processing. This finding suggests that other factors beyond expertise in individuation may be involved in holistic processing.

One way we can better understand holistic processing is by researching the neural mechanisms that underlie this process. For faces, several studies have suggested a particular importance of the fusiform face area (FFA). FFA responds more to whole faces than to the same faces with the facial parts scrambled (Brandman & Yovel, 2016; Liu, Harris, & Kanwisher, 2010; Zhao et al., 2014). Furthermore, neural activity in FFA has been shown to correspond with holistic processing of composite faces (Andrews, Davies-Thompson, Kingstone, & Young, 2010; Goffaux, Schiltz, Mur, & Goebel, 2013; Schiltz, Dricot, Goebel, & Rossion, 2010; Schiltz & Rossion, 2006), though it has also been suggested this area contains both holistic and part-based representations of faces (Harris & Aguirre, 2008, 2010). A similar involvement of the FFA has been shown for processing of expertise objects. FFA has been shown to be activated in bird experts viewing bird images, and car experts viewing car images (Gauthier, Skudlarski, Gore, & Anderson, 2000; Xu, 2005), and activity in the FFA has been found to increase with expertise in recognising novel objects (Gauthier, Tarr, Anderson, Skudlarski, & Gore, 1999). Furthermore, evidence of holistic processing of cars in the right FFA2 (a subsection of the FFA) has been shown to correlate with the level of car expertise (Ross et al., 2018). These studies demonstrate an importance of the FFA in holistic processing of faces and objects of expertise. However, it is not yet clear if the FFA is also involved in holistic processing of non-expertise objects.

In this study, we investigated the neural mechanisms underlying holistic processing of non-expertise line pattern objects. We recorded the brain activity of participants using functional magnetic resonance imaging (fMRI) as they performed a composite task with line pattern objects. This task involved participants viewing pairs of composite line patterns, one shown after the other, and making top-same or top-different judgements as to the identity of the top-half of the line pattern. This task has previously been shown to elicit evidence of holistic processing of these line pattern objects in behaviour (Zhao et al., 2016). We investigated two components of the composite paradigm. Firstly, we investigated if there

were any overall activation differences between aligned and misaligned line patterns, as aligned line patterns are thought to be processed more holistically than misaligned line patterns. Secondly, we investigated if any regions showed activity related to the change in the perception of the identity of the top-half of the line pattern that is induced by the inability of participants to ignore the bottom half of the aligned line patterns.

We investigated changes in brain activity during the composite task in a variety of brain regions that we hypothesized might be involved in holistic processing of the line pattern objects. First, we localized a region in the lateral occipital cortex (LOC) that is known to be involved in object processing (Malach et al., 1995) and has been shown to respond to local contour elements that are perceptually grouped into a global shape (Altmann, Bülthoff, & Kourtzi, 2003). Second, we localized three scene-responsive regions, the transverse occipital sulcus (TOS, also known as OPA), the retrosplenial cortex (RSC) and the parahippocampal place area (PPA). The PPA and TOS/OPA have been shown to respond to a holistically processed scene stimulus compared to a control stimulus matched in low-level visual content (Schindler & Bartels, 2016), and the PPA and RSC have been shown to respond more to intact scenes as compared to fractured ones (Kamps, Julian, Kubilius, Kanwisher, & Dilks, 2016). These studies suggest that these scene-responsive regions are involved in holistically grouping objects together into a scene, which we hypothesized could also relate to grouping of the local elements of the line pattern objects into a holistic representation. Third, we localized regions responsive to faces, the FFA that has previously been found to be involved in holistic processing of faces and objects of expertise (Ross et al., 2018; Schiltz et al., 2010; Schiltz & Rossion, 2006), as well as the occipital face area (OFA) and anterior temporal face area (ATFA). Fourth, we localized a region in the superior parietal lobe (SPL) that has been shown to be activated during holistic processing of several bi-stable Gestalt stimuli (Grassi, Zaretskaya, & Bartels, 2016, 2018; Zaretskaya, Anstis, & Bartels, 2013). As Gestalt information has been shown to be important for holistic processing of line pattern objects (Zhao et al., 2016), we hypothesized the SPL might be involved in holistic processing of these line pattern objects.

6.2. Materials and methods

6.2.1. Participants

Twenty-two participants (16 female, 20-39 years old) took part in the study. One participant was excluded from the fMRI data analyses due to excessive head movement during scanning. All participants provided written informed consent prior to their participation in the study. The experimental procedure was approved by the ethics committee of the University Clinic Tübingen.

6.2.2. Stimuli

Our experimental stimuli were created from twelve pairs of line patterns selected from a dataset used in a previous study (Zhao et al., 2016). These line pattern stimuli were shown to successfully elicit evidence of holistic processing. The twelve line pattern pairs were separated into top and bottom halves, and within each pair the top and bottom halves of the patterns were recombined (e.g. top half of line pattern A and bottom half of line pattern B) to create composite line patterns. Although the top and bottoms halves of each pair were different, they could be recombined with no disruption of the line continuity between the top and bottom halves. Line pattern stimuli had an average height of 3.7° and width of 3.6° of visual angle. We additionally created misaligned line pattern stimuli by shifting the bottom halves of the line patterns 1.0° of visual angle to the left. This misalignment has previously been shown to disrupt holistic processing (Zhao et al., 2016). We showed a horizontal black line (0.03° of visual angle) between the top and bottom halves of each line pattern, so that participants could clearly separate the two halves. We created eight additional pairs of line pattern stimuli, using the same method, which were used for practice trials prior to the main experiment.

6.2.3. Experimental procedure

Participants lay supine in the scanner and viewed the stimuli on a screen positioned behind their head, via a mirror attached to the head coil. The screen was positioned 82 cm from the participant, and spanned 28° x 16° of visual angle in horizontal and vertical

directions respectively. Stimuli were presented via a projector with resolution 1920x1080. The experiment was programmed with Matlab 2013b using the Psychophysics Toolbox extensions (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007) on a Windows PC.

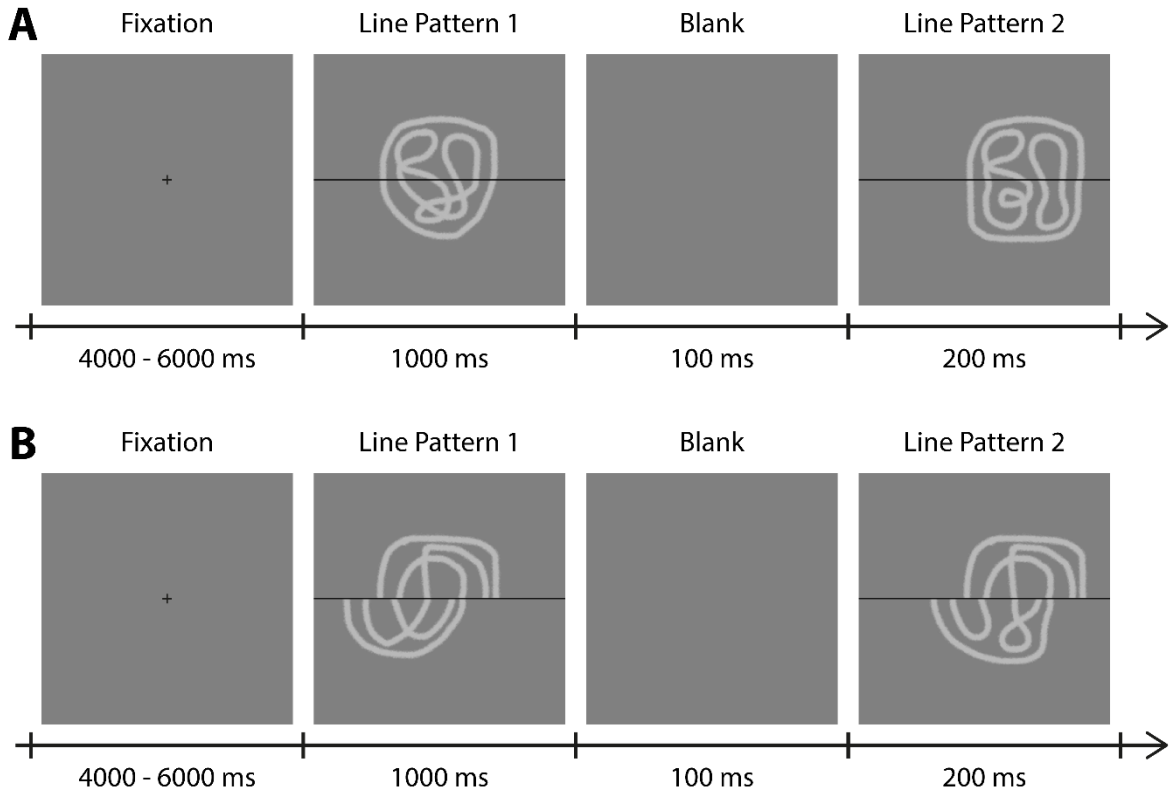

Figure 1. Illustration of the experiment trial procedure. Participants fixated for either 4 or 6 s, then viewed a first line pattern, followed by a blank screen and then a second line pattern. Participants then responded during the next fixation whether the top-halves of the two line patterns were the same or different. (A) illustrates the procedure for the aligned line pattern stimuli, and (B) illustrates the procedure for the misaligned line pattern stimuli.

Participants performed a composite task while we recorded their brain activity using fMRI. Figure 1 illustrates the trial procedure of the composite task. Participants viewed a central fixation cross for 4 s or 6 s (50% of trials each, random order). The first line pattern was shown centrally on the screen for 1 s, followed by a blank screen (presented for 100 ms), then the second line pattern was shown, 1.2° of visual angle offset to the right of the centre of the screen, for 200 ms. Participants responded during the next fixation using a

button press whether they judged the top halves of the two line patterns to be the same or different. They were instructed beforehand to ignore the bottom halves of the line patterns. We counterbalanced which fingers participants used to respond top-same/top-different across participants.

The experimental design consisted of 8 conditions of a 2 x 2 x 2 factorial design (see Fig. 2). The factors were *alignment*, whether the bottom halves of the line patterns were aligned or misaligned with the top halves, *congruency*, whether the bottom half of the second line pattern was congruent with respect to the top half of the second line pattern or not (e.g. congruent trials are when the bottom-half is the same if the top-half is the same and the bottom-half is different if the top-half is different) and *same/different*, whether the top halves of the two line patterns were the same or different from each other. Each participant completed 3 runs, where each run contained 64 trials (8 repetitions per condition). Conditions were presented in a carryover counterbalanced design, such that each condition was preceded by every other condition once per run (Brooks, 2012). This was to avoid biases from carryover blood-oxygen-level dependent (BOLD) activation from a previous condition (Aguirre, 2007).

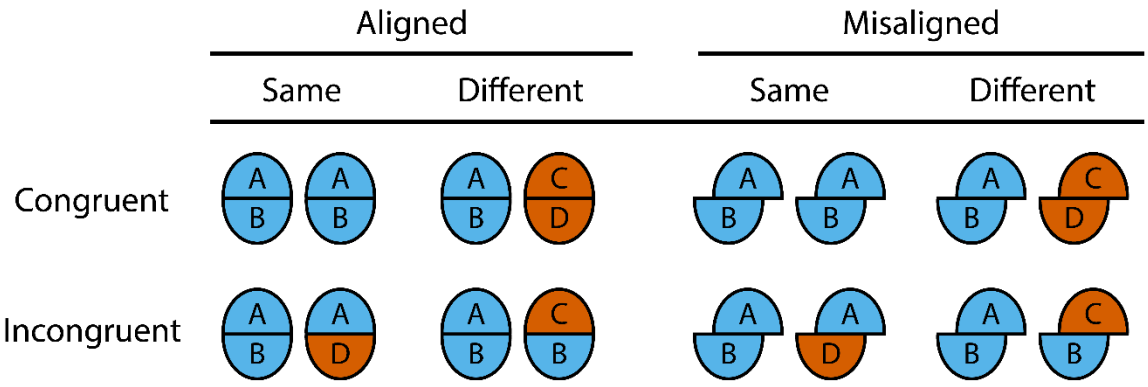

Figure 2. Experimental conditions. The conditions consisted of a 2 x 2 x 2 factorial design, with factors *alignment*, whether the top and bottom halves of the line patterns were aligned or misaligned, *same or different*, whether the top halves of the line patterns were the same or different from each other and *congruency*, whether the bottom half of line pattern 2 was congruent with respect to the top half of line pattern 2 or not (e.g. *congruent* trials are when the bottom-half is the same if the top-half is the same and the bottom-half is different if the top-half is different).

6.2.4. Localizer stimuli and procedure

Participants completed 2 runs of a localizer experiment, which we used to define object-, scene- and face-responsive ROIs. Participants viewed blocks containing grayscale objects, scenes, faces and phase-scrambled scenes. Phase-scrambled scenes were Fourier-scrambled versions of the original scene images. As the object and face stimuli were smaller in size than the scenes and phase-scrambled scenes, we showed the phase-scrambled scenes as background images to the objects and faces in order to keep the visual field size of the stimuli constant in all blocks. Blocks were presented in a carryover counterbalanced sequence (Brooks, 2012). In each block 8 images were shown, where each image was shown for 1.8 s, followed by a 0.2 s blank, grey screen. Participants performed a one-back task on the images (with repetitions once every 9 s on average) to keep their attention to the stimuli.

6.2.5. Imaging parameters

Images were acquired using a 3T Siemens Prisma scanner with a 64-channel head coil (Siemens, Erlangen, Germany). Functional T2* echoplanar images (EPI) were acquired using a sequence with the following parameters; multiband acceleration factor 2, TR 1.39 s, TE 30 ms, flip angle 68°, FOV 192x192 mm. Volumes consisted of 42 slices, with an isotropic voxel size of 3x3x3 mm. The first 8 volumes of each run were discarded to allow for equilibration of the T1 signal. For each participant a high-resolution T1-weighted anatomical scan was acquired with the following parameters; TR 2 s, TE 3.06 ms, FOV 232x256 mm, 192 slices, isotropic voxel size of 1x1x1 mm.

6.2.6. fMRI data preprocessing

fMRI data was preprocessed with SPM12 (http://www.fil.ion.ucl.ac.uk/spm/). Functional images were slice-time corrected, realigned and coregistered to the anatomical image. The images were then normalized to MNI (Montreal Neurological Institute) space and spatially smoothed with a 6 mm full-width at half-maximum Gaussian kernel.

6.2.7. Definition of regions of interest

Table 1 shows the average locations and sizes of our ROIs. We used fMRI data from the localizer runs to define object-, scene- and face-responsive ROIs. The object-responsive LOC was defined using the contrast objects > phase-scrambled scenes (Malach et al., 1995). Scene-responsive ROIs were defined using the contrast scenes > faces, and these ROIs included TOS/OPA, RSC and PPA (Epstein & Kanwisher, 1998; Grill-Spector, 2003; Maguire, 2001). The face-responsive OFA, FFA and ATFA were defined using the contrast faces > objects (Gauthier, Tarr, et al., 2000; Kanwisher, McDermott, & Chun, 1997; Rajimehr, Young, & Tootell, 2009; Tsao, Moeller, & Freiwald, 2008). For each contrast we selected a threshold of $p < 0.001$ uncorrected to define active voxels. We defined each ROI individually for each participant, and selected all active voxels within a sphere (radius 6 mm) centred on the peak of activity for each ROI.

We defined ROIs SPL and V1 based on their anatomical location for each participant, and using the contrast: activity during stimulus presentation > fixation interval between trials. This contrast was orthogonal to the activity differences investigated in our results (Friston, Rotshtein, Geng, Sterzer, & Henson, 2006). We selected active voxels using a $p < 0.05$ familywise error rate (FWE) corrected threshold. To define SPL, we selected all active voxels within a sphere (radius 6 mm) centred on the peak of activity in the SPL. To define V1, we first defined the entirety of V1 using anatomical labels generated by Freesurfer (Hinds et al., 2009) (https://surfer.nmr.mgh.harvard.edu/), and then selected the posterior, foveal part of V1 that was activated by the contrast as our V1 ROI.

Table 1

Average MNI coordinates and volume of each ROI (± standard deviations). N indicates the number of participants each ROI was identified in.

ROI	hem	X	y	z	Volume (mm^3)	N
LOC	left	-42 ± 4.7	-80 ± 5.8	-4 ± 5.1	221 ± 14.1	21
	right	44 ± 3.8	-80 ± 5.1	-5 ± 5.8	212 ± 27.6	21
TOS/OPA	left	-31 ± 5.7	-86 ± 4.3	21 ± 7.5	212 ± 25.2	21
	right	36 ± 5.4	-82 ± 4.0	20 ± 6.8	219 ± 19.5	21
RSC	left	-17 ± 3.2	-59 ± 3.2	14 ± 3.3	183 ± 65.3	20
	right	20 ± 3.2	-57 ± 4.6	17 ± 4.5	199 ± 53.5	20
PPA	left	-28 ± 2.9	-46 ± 4.1	-10 ± 4.2	221 ± 11.9	21
	right	29 ± 3.2	-46 ± 6.4	-10 ± 3.4	221 ± 16.6	20
OFA	left	-38 ± 4.3	-78 ± 7.1	-12 ± 3.2	188 ± 52.4	21
	right	41 ± 3.4	-77 ± 6.7	-12 ± 5.4	201 ± 50.3	21
FFA	left	-41 ± 3.9	-52 ± 5.5	-19 ± 4.1	175 ± 58.5	21
	right	42 ± 3.1	-50 ± 5.2	-19 ± 3.9	203 ± 37.1	21
ATFA	left	-35 ± 5.9	-7 ± 6.8	-35 ± 6.3	78 ± 47.7	11
	right	34 ± 3.8	-4 ± 3.6	-39 ± 4.4	122 ± 55.8	15
SPL	left	-26 ± 4.7	-58 ± 5.0	51 ± 7.2	217 ± 19.6	21
	right	26 ± 4.1	-56 ± 5.1	51 ± 6.3	219 ± 21.8	21
V1	left	-12 ± 3.1	-98 ± 2.5	-7 ± 4.0	962 ± 319.0	21
	right	14 ± 3.3	-96 ± 2.0	-5 ± 3.9	979 ± 249.1	21

6.2.8. Behavioural analyses

We calculated participants' behavioural responses during the composite task using accuracy (% correct) and reaction times. For each behavioural measure we first performed a 2 (alignment) x 2 (congruency) x 2 (top-same/top-different) repeated measures ANOVA to investigate if there was a triple interaction between the three factors. We then performed further behavioural analyses separately for top-same and top-different conditions as previous studies have shown that evidence of holistic processing is different for these conditions (Goffaux, 2012; Goffaux et al., 2013). We performed 2 (alignment) x 2 (congruency) repeated measures ANOVAs to investigate if there was behavioural evidence of holistic processing during the top-same and top-different conditions. For evidence of holistic processing we expected to find an effect of congruency and an interaction between congruency and alignment, driven by a difference in behavioural responses between congruent and incongruent conditions that is greater for the aligned conditions compared to the misaligned ones. We confirmed this pattern of results using follow up *t*-tests. For accuracy, we expected a lower accuracy of participants for incongruent aligned conditions compared to congruent aligned conditions, driven by participants change in perception of the top-half of the line pattern caused by holistic processing. For reaction times, we expected participants to respond slower on incongruent aligned conditions compared to congruent aligned conditions, due to participants taking longer to make a response decision.

6.2.9. fMRI analyses

We modelled a GLM using SPM12 for each participant, containing regressors for the 8 conditions and the 6 realignment regressors from the motion correction during preprocessing. We excluded any trial where the participant did not make a response to indicate a top-same/top-different judgment from the 8 condition regressors. We performed all fMRI analyses in our 9 ROIs as well as in whole-brain analyses. ROI analyses were Bonferroni-corrected for N = 9 ROIs (*p* values are shown uncorrected unless stated as Bonferroni-corrected) and whole-brain analyses were False Discovery Rate (FDR) corrected. We first performed a 2 (alignment) x 2 (congruency) x 2 (top-same/top-different) repeated measures ANOVA to investigate if any regions showed a triple interaction between the

three factors. We then performed analyses between specific conditions to investigate different aspects of the composite effect.

First, we investigated if any regions showed differences in neural activity between aligned and misaligned congruent conditions. As aligned patterns are thought to be processed more holistically than misaligned ones (Zhao et al., 2016), we hypothesized that regions involved with holistic processing of these line patterns might show differences in the level of activation during these conditions. We included only congruent conditions in these analyses, as for incongruent conditions there is a change in the participants' perception of the line pattern in the aligned trials, but not in misaligned trials. Therefore, any differences in neural activation between aligned and misaligned incongruent trials could be due to the change in perception of the line pattern rather than due to differences in processing of the aligned and misaligned trials. We performed a 2 (alignment) x 2 (top-same/top-different) repeated measures ANOVA to test whether any regions showed a main effect of alignment.

Second, we investigated if any regions showed differences in neural activity related to the change in the perception of the top-half of the line pattern. We performed this analysis separately for top-same and top-different conditions as previous studies have found differences in the strength of evidence of holistic processing for these conditions (Goffaux, 2012; Goffaux et al., 2013). For top-same and top-different conditions we performed 2 (alignment) x 2 (congruency) repeated measures ANOVAs to test whether any regions showed a main effect of congruency and an interaction between congruency and alignment. For evidence of holistic processing, we expected to find a difference in neural responses between congruent and incongruent conditions that is greater for the aligned conditions compared to the misaligned ones. We performed follow-up *t*-tests in any significant regions to confirm this pattern of responses.

6.3. Results

6.3.1. Behavioural results

We measured participants' behavioural responses to the 8 conditions of the composite task using accuracy (% correct) and reaction times. 2 (alignment) x 2 (congruency) x 2 (top-same/top-different) repeated measures ANOVAs revealed significant triple-interactions for both accuracy ($F_{1,20}$ = 47.69, p = 1.0 × 10^{-6}, η_p^2 = 0.70) and reaction times ($F_{1,20}$ = 10.11, p = 0.0047, η_p^2 = 0.34). We then conducted further analyses separately for top-same and top-different conditions to investigate which conditions show differences related to holistic processing of the line pattern objects.

6.3.1.1. Accuracy and reaction times for top-same conditions

Figure 3 shows accuracy and reaction times for the top-same conditions of the composite task. 2 (alignment) x 2 (congruency) repeated measures ANOVAs revealed both significant effects of congruency (accuracy: $F_{1,20}$ = 56.19, p = 3.1 × 10^{-7}, η_p^2 = 0.74; reaction times: $F_{1,20}$ = 14.12, p = 0.0012, η_p^2 = 0.41) and significant interactions between congruency and alignment (accuracy: $F_{1,20}$ = 58.31, p = 2.4 × 10^{-7}, η_p^2 = 0.74; reaction times: $F_{1,20}$ = 23.42, p = 1.0 × 10^{-4}, η_p^2 = 0.54) for both behavioural measures. Follow-up paired t-tests confirmed that the congruency effect was significant for the aligned conditions (accuracy: M = 37.50 %, SE = 4.68 %; t_{20} = 8.01, p = 1.1 × 10^{-7}, Cohen's d_z = 1.75; reaction times: M = 0.17 s, SE = 0.038 s; t_{20} = 4.55, p = 1.9 × 10^{-4}, Cohen's d_z = 0.99) but not for the misaligned conditions (accuracy: M = 3.37 %, SE = 1.72 %; t_{20} = 1.97, p = 0.063, Cohen's d_z = 0.43; reaction times: M = -0.01 s, SE = 0.014 s; t_{20} = -0.85, p = 0.40, Cohen's d_z = -0.19). Thus, these behavioural results show evidence of holistic processing during the top-same conditions of the composite task.

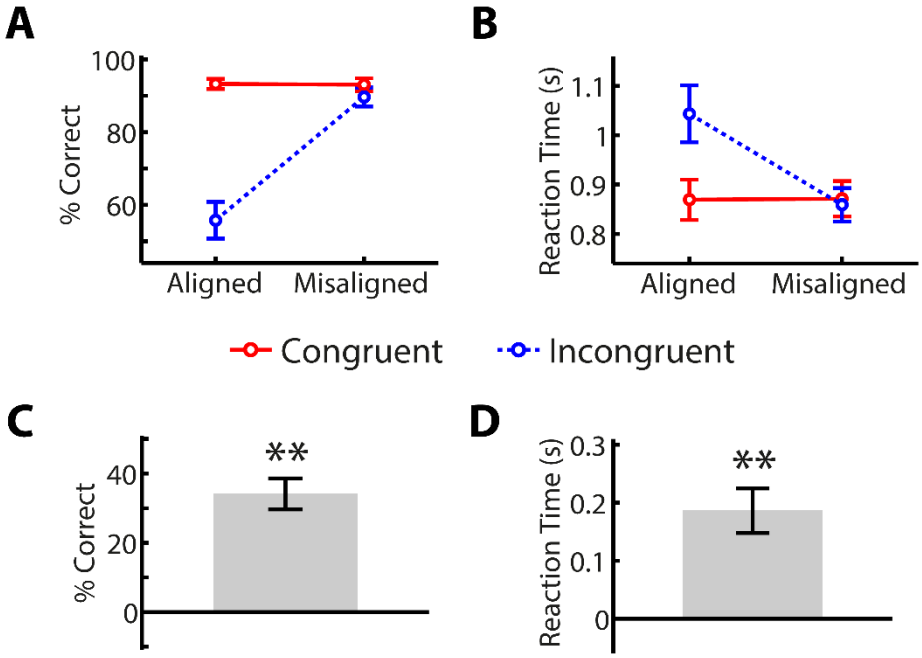

Figure 3. Behavioural results for the top-same conditions of the composite task. (A) shows accuracy (% correct) as a function of alignment and congruency, and (B) shows reaction times as a function of alignment and congruency. (C) and (D) show the interaction effect between alignment and congruency for accuracy (C) and reaction times (D). Error bars indicate ±1 SEM. ** indicates $p < 0.001$.

6.3.1.2. Accuracy and reaction times for top-different conditions

Figure 4 shows accuracy and reaction times for the top-different conditions of the composite task. For reaction times, a 2 (alignment) x 2 (congruency) repeated measures ANOVA revealed a significant effect of congruency ($F_{1,20} = 4.47$, $p = 0.047$, $\eta_p^2 = 0.18$) and significant interaction between congruency and alignment ($F_{1,20} = 6.67$, $p = 0.018$, $\eta_p^2 = 0.25$). For accuracy, a 2 (alignment) x 2 (congruency) repeated measures ANOVA revealed a significant effect of congruency ($F_{1,20} = 4.42$, $p = 0.048$, $\eta_p^2 = 0.18$) but no significant interaction between congruency and alignment ($F_{1,20} = 0.36$, $p = 0.55$, $\eta_p^2 = 0.018$). Follow-up paired *t*-tests investigating the reaction time effects confirmed that the congruency effect was significant for the aligned conditions ($M = 0.044$ s, $SE = 0.013$ s; $t_{20} = 3.37$, $p = 0.0030$, Cohen's $d_z = 0.74$) but not for the misaligned conditions ($M = -0.0075$ s, $SE = 0.013$ s;

t_{20} = -0.56, p = 0.58, Cohen's d_z = -0.12). These reaction time results show evidence of holistic processing during the top-different conditions of the composite task.

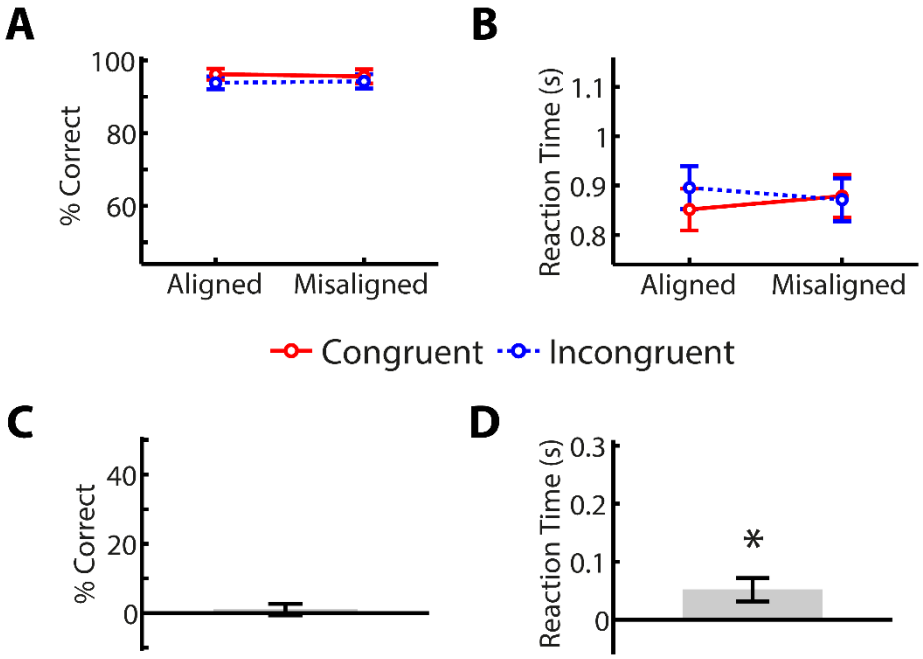

Figure 4. Behavioural results for the top-different conditions of the composite task. (A) shows accuracy (% correct) as a function of alignment and congruency, and (B) shows reaction times as a function of alignment and congruency. (C) and (D) show the interaction effect between alignment and congruency for accuracy (C) and reaction times (D). Error bars indicate ±1 SEM. * indicates p < 0.05.

6.3.2. fMRI results

We first tested whether any of our ROIs showed a triple interaction effect between alignment, congruency and top-same/top-different conditions using 2 x 2 x 2 repeated measures ANOVAs. We did not identify any regions showing a significant triple interaction effect in these analyses (LOC: $F_{1,20}$ = 1.51, p = 0.23, η_p^2 = 0.070; TOS/OPA: $F_{1,20}$ = 0.020, p = 0.89, η_p^2 = 0.0010; RSC: $F_{1,19}$ = 0.23, p = 0.64, η_p^2 = 0.012; PPA: $F_{1,20}$ = 0.076, p = 0.79, η_p^2 = 0.0038; OFA: $F_{1,20}$ = 1.63, p = 0.22, η_p^2 = 0.075; FFA: $F_{1,20}$ = 0.40, p = 0.54, η_p^2 = 0.019; ATFA:

$F_{1,14}$ = 0.57, p = 0.46, η_p^2 = 0.039; SPL: $F_{1,20}$ = 4.12, p = 0.056, η_p^2 = 0.017; V1: $F_{1,20}$ = 0.89, p = 0.36, η_p^2 = 0.043).

Next, we tested for differences between specific conditions in order to investigate which regions show neural activity related to specific aspects of the composite task. First, we investigated which regions show differences in neural responses to aligned and misaligned line pattern objects. Aligned line patterns are thought to be processed more holistically than misaligned line patterns, thus we predicted that regions involved in holistic processing might show differences in neural activity between these conditions. Second, we investigated which regions show differences in neural activity related to the change in the perception of the top-half of the line pattern object that is induced by the inability to ignore the aligned bottom-half of the line pattern object. Here, we investigated which regions show a difference in neural activity between congruent and incongruent conditions that is greater for aligned as compared to misaligned line pattern stimuli. We performed these analyses separately for top-same and top-different conditions due to the known differences in the strength of the behavioural evidence of holistic processing for top-same and top-different conditions (Goffaux, 2012; Goffaux et al., 2013).

6.3.2.1. Aligned vs. misaligned conditions

We investigated which brain regions show differences in neural responses to aligned versus misaligned congruent conditions (Fig. 5). We included only the congruent conditions in these analyses as in the incongruent conditions there is also a change in the perception of the top-half of the line pattern that differs between aligned and misaligned conditions. We performed 2 (alignment) x 2 (top-same/top-different) repeated measures ANOVAs to investigate which regions show a main effect of alignment for the congruent conditions. For object- and scene-responsive ROIs (Fig. 5A) we did not find a significant effect of alignment in any region (LOC: $F_{1,20}$ = 2.29, p = 0.15, η_p^2 = 0.10; TOS/OPA: $F_{1,20}$ = 0.065, p = 0.80, η_p^2 = 0.0032; RSC: $F_{1,19}$ = 0.13, p = 0.72, η_p^2 = 0.0068; PPA: $F_{1,20}$ = 2.71, p = 0.12, η_p^2 = 0.12). For all other ROIs (Fig. 5B), we found a marginally significant effect of alignment in the SPL ($F_{1,20}$ = 5.10, p = 0.035, η_p^2 = 0.20), but this result did not survive Bonferroni-correction for N = 9 ROIs. We did not find a significant effect of alignment in any of the other ROIs tested (OFA: $F_{1,20}$ = 1.55, p = 0.23, η_p^2 = 0.072; FFA: $F_{1,20}$ = 2.26, p = 0.15, η_p^2 = 0.10; ATFA: $F_{1,14}$ = 0.016, p

= 0.90, η_p^2 = 0.0011; V1: $F_{1,20}$ = 0.012, p = 0.91, η_p^2 = 6.1 x 10^{-4}). We performed a whole-brain analysis to investigate if any regions outside of our ROIs showed significant differences in activation between aligned and misaligned stimuli. We did not identify any regions in this analysis.

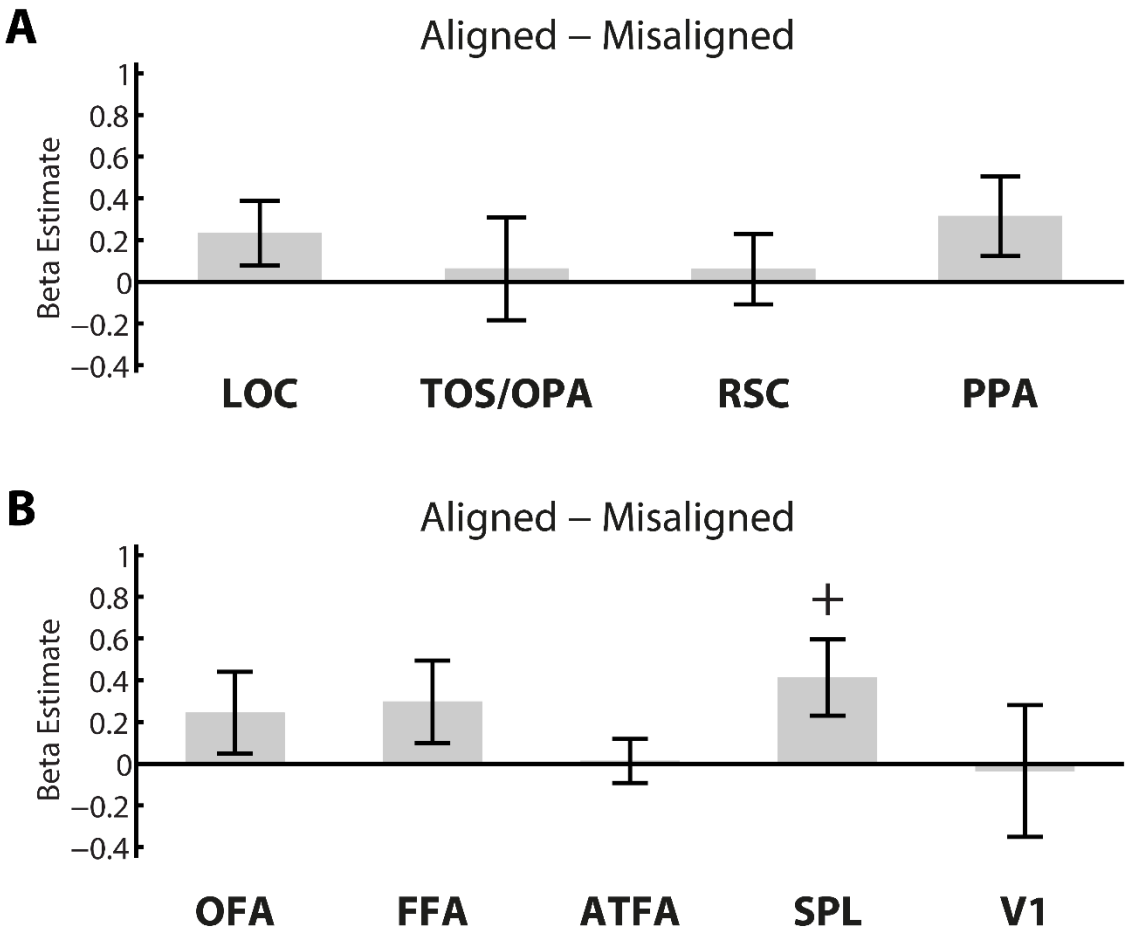

Figure 5. Differences in neural responses between aligned and misaligned congruent conditions, using the contrast aligned – misaligned. (A) shows results for object- and scene-responsive ROIs, (B) shows results for face-responsive, perceptual-grouping related and early visual ROIs. Error bars indicate ±1 SEM. + indicates $p < 0.05$ uncorrected.

6.3.2.2. Congruency & alignment interaction for top-same conditions

We investigated if any brain regions showed a difference in neural activity between the congruent and incongruent top-same conditions of the composite task, and whether this difference was greater for the aligned conditions compared to the misaligned conditions. This pattern of responses would show neural activity consistent with the change in the perception of the top-half of the line pattern that is induced by holistic processing.

For the object- and scene-responsive ROIs (Fig. 6), 2 (congruency) x 2 (alignment) repeated measures ANOVAs did not identify any significant main effects of congruency (LOC: $F_{1,20} = 0.059$, $p = 0.81$, $\eta_p^2 = 0.0030$; TOS/OPA: $F_{1,20} = 0.12$, $p = 0.73$, $\eta_p^2 = 0.0060$; RSC: $F_{1,19} = 4.21$, $p = 0.054$, $\eta_p^2 = 0.18$; PPA: $F_{1,20} = 0.87$, $p = 0.36$, $\eta_p^2 = 0.042$) or significant interactions between congruency and alignment (LOC: $F_{1,20} = 0.73$, $p = 0.40$, $\eta_p^2 = 0.035$; TOS/OPA: $F_{1,20} = 0.054$, $p = 0.82$, $\eta_p^2 = 0.0027$; RSC: $F_{1,19} = 3.0 \times 10^{-4}$, $p = 0.99$, $\eta_p^2 = 1.6 \times 10^{-5}$; PPA: $F_{1,20} = 0.34$, $p = 0.57$, $\eta_p^2 = 0.017$) for the top-same conditions. We next tested whether any of our other ROIs would show an effect of congruency or interaction between congruency and alignment (Fig. 7). 2 x 2 repeated measures ANOVAs in these ROIs did not identify any significant effects of congruency (OFA: $F_{1,20} = 0.040$, $p = 0.84$, $\eta_p^2 = 0.0020$; FFA: $F_{1,20} = 0.018$, $p = 0.90$, $\eta_p^2 = 8.9 \times 10^{-4}$; ATFA: $F_{1,14} = 3.92$, $p = 0.068$, $\eta_p^2 = 0.22$; SPL: $F_{1,20} = 2.99$, $p = 0.099$, $\eta_p^2 = 0.13$; V1: $F_{1,20} = 0.0053$, $p = 0.94$, $\eta_p^2 = 2.7 \times 10^{-4}$), or significant interactions between congruency and alignment (OFA: $F_{1,20} = 0.44$, $p = 0.51$, $\eta_p^2 = 0.022$; FFA: $F_{1,20} = 0.66$, $p = 0.43$, $\eta_p^2 = 0.032$; ATFA: $F_{1,14} = 2.57$, $p = 0.13$, $\eta_p^2 = 0.16$; SPL: $F_{1,20} = 2.85$, $p = 0.11$, $\eta_p^2 = 0.12$; V1: $F_{1,20} = 0.82$, $p = 0.38$, $\eta_p^2 = 0.040$). We performed a whole-brain analysis to investigate if any other brain regions would show a significant interaction between congruency and alignment in the top-same conditions. We did not identify any regions in this analysis.

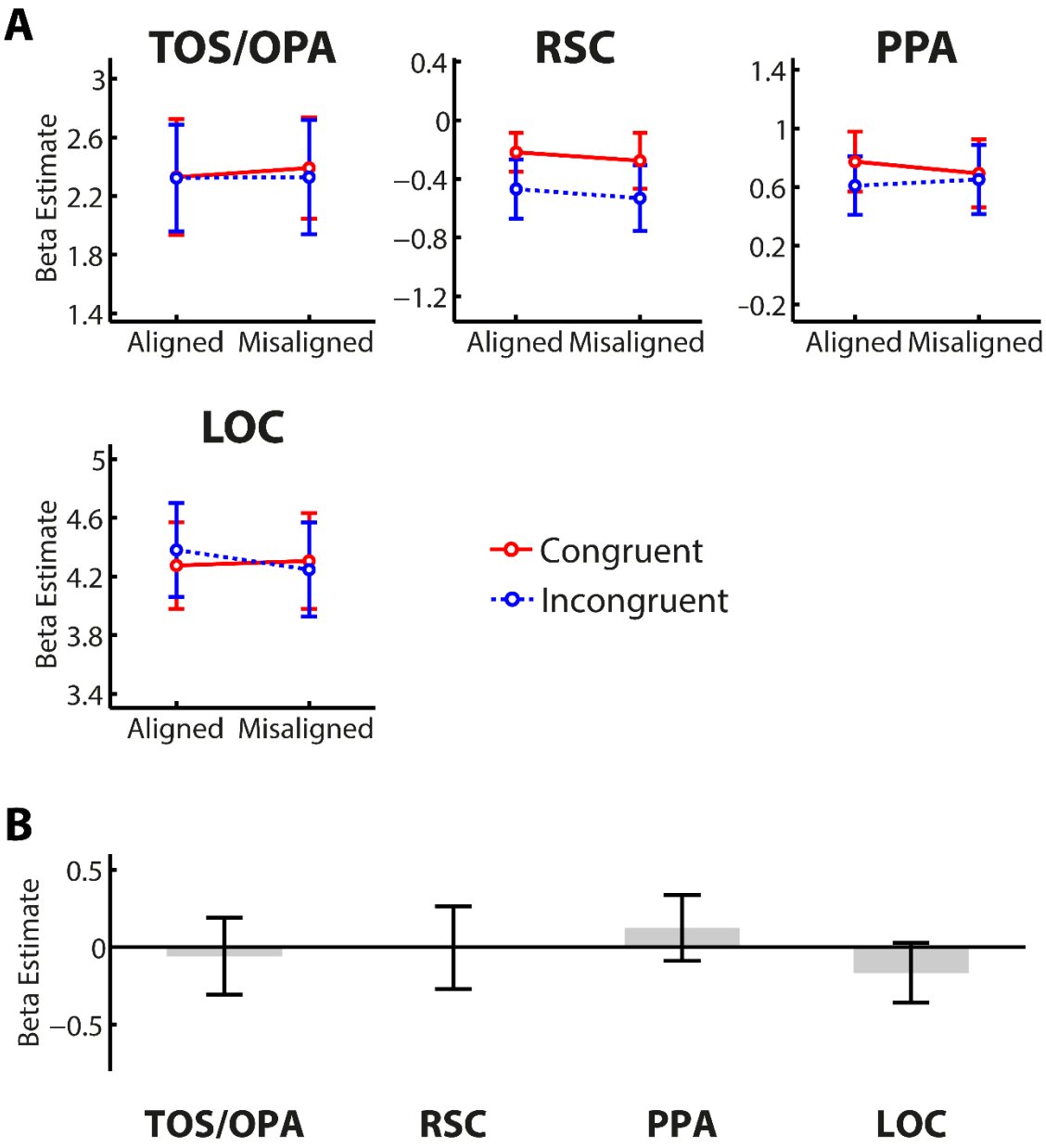

Figure 6. Neural responses to the top-same conditions of the composite task in object- and scene-responsive ROIs. (A) shows neural responses as a function of alignment and congruency. (B) shows the interaction effect between alignment and congruency. Error bars indicate ±1 SEM.

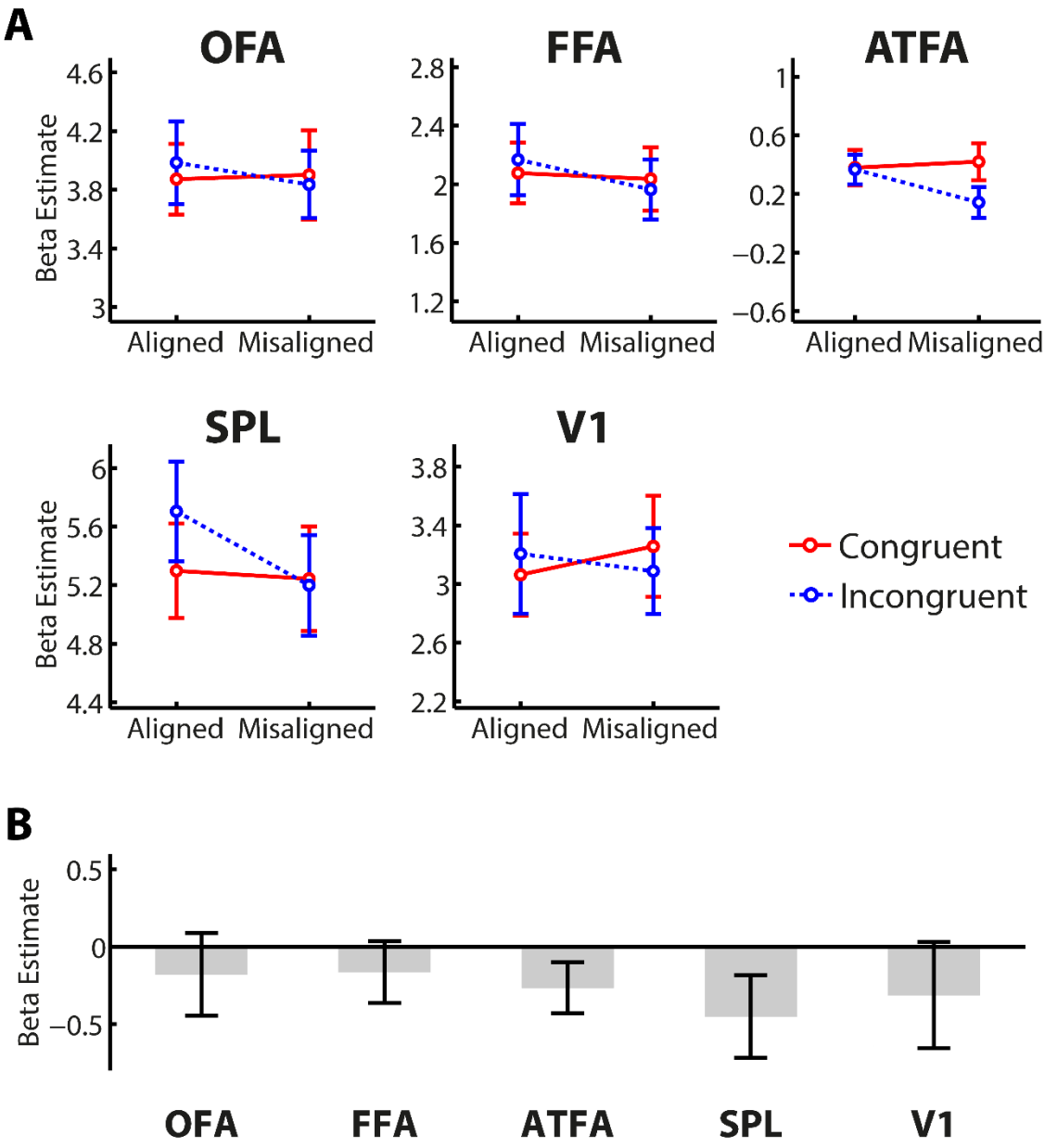

Figure 7. Neural responses to the top-same conditions of the composite task in face-responsive, perceptual-grouping related and early visual ROIs. (A) shows neural responses as a function of alignment and congruency. (B) shows the interaction effect between alignment and congruency. Error bars indicate ±1 SEM.

6.3.2.3. Congruency & alignment interaction for top-different conditions

We investigated if any brain regions showed a difference in neural activity between the congruent and incongruent top-different conditions of the composite task, and whether this difference was greater for the aligned compared to the misaligned conditions. As previously, this pattern of responses would suggest neural responses consistent with the change in the perception of the top-half of the line pattern that is induced by holistic processing during the aligned incongruent conditions.

For object- and scene-responsive ROIs (Fig. 8), 2 (alignment) x 2 (congruency) repeated measures ANOVAs did not identify any significant main effects of congruency (LOC: $F_{1,20} = 0.58$, $p = 0.45$, $\eta_p^2 = 0.028$; TOS/OPA: $F_{1,20} = 0.66$, $p = 0.43$, $\eta_p^2 = 0.032$; RSC: $F_{1,19} = 0.17$, $p = 0.68$, $\eta_p^2 = 0.0089$; PPA: $F_{1,20} = 0.084$, $p = 0.78$, $\eta_p^2 = 0.0042$) or significant interactions between congruency and alignment (LOC: $F_{1,20} = 0.59$, $p = 0.45$, $\eta_p^2 = 0.029$; TOS/OPA: $F_{1,20} = 0.30$, $p = 0.59$, $\eta_p^2 = 0.015$; RSC: $F_{1,19} = 0.079$, $p = 0.39$, $\eta_p^2 = 0.040$; PPA: $F_{1,20} = 0.033$, $p = 0.86$, $\eta_p^2 = 0.0017$) for the top-different conditions. We further tested whether any of our other ROIs would show an effect of congruency or interaction between congruency and alignment for the top-different conditions (Fig. 9). We did not identify any regions that showed a significant effect of congruency (OFA: $F_{1,20} = 0.46$, $p = 0.50$, $\eta_p^2 = 0.023$; FFA: $F_{1,20} = 0.37$, $p = 0.55$, $\eta_p^2 = 0.018$; ATFA: $F_{1,14} = 0.15$, $p = 0.70$, $\eta_p^2 = 0.011$; SPL: $F_{1,20} = 0.055$, $p = 0.82$, $\eta_p^2 = 0.0027$; V1: $F_{1,20} = 0.60$, $p = 0.45$, $\eta_p^2 = 0.029$) or significant interaction between congruency and alignment (OFA: $F_{1,20} = 1.40$, $p = 0.25$, $\eta_p^2 = 0.065$; FFA: $F_{1,20} = 0.012$, $p = 0.91$, $\eta_p^2 = 6.0 \times 10^{-4}$; ATFA: $F_{1,14} = 0.60$, $p = 0.45$, $\eta_p^2 = 0.04$; SPL: $F_{1,20} = 0.30$, $p = 0.59$, $\eta_p^2 = 0.015$; V1: $F_{1,20} = 0.25$, $p = 0.62$, $\eta_p^2 = 0.012$). We performed a whole-brain analysis to investigate if any regions outside of our ROIs would show an interaction between congruency and alignment for the top-different conditions. We did not identify any regions in this analysis.

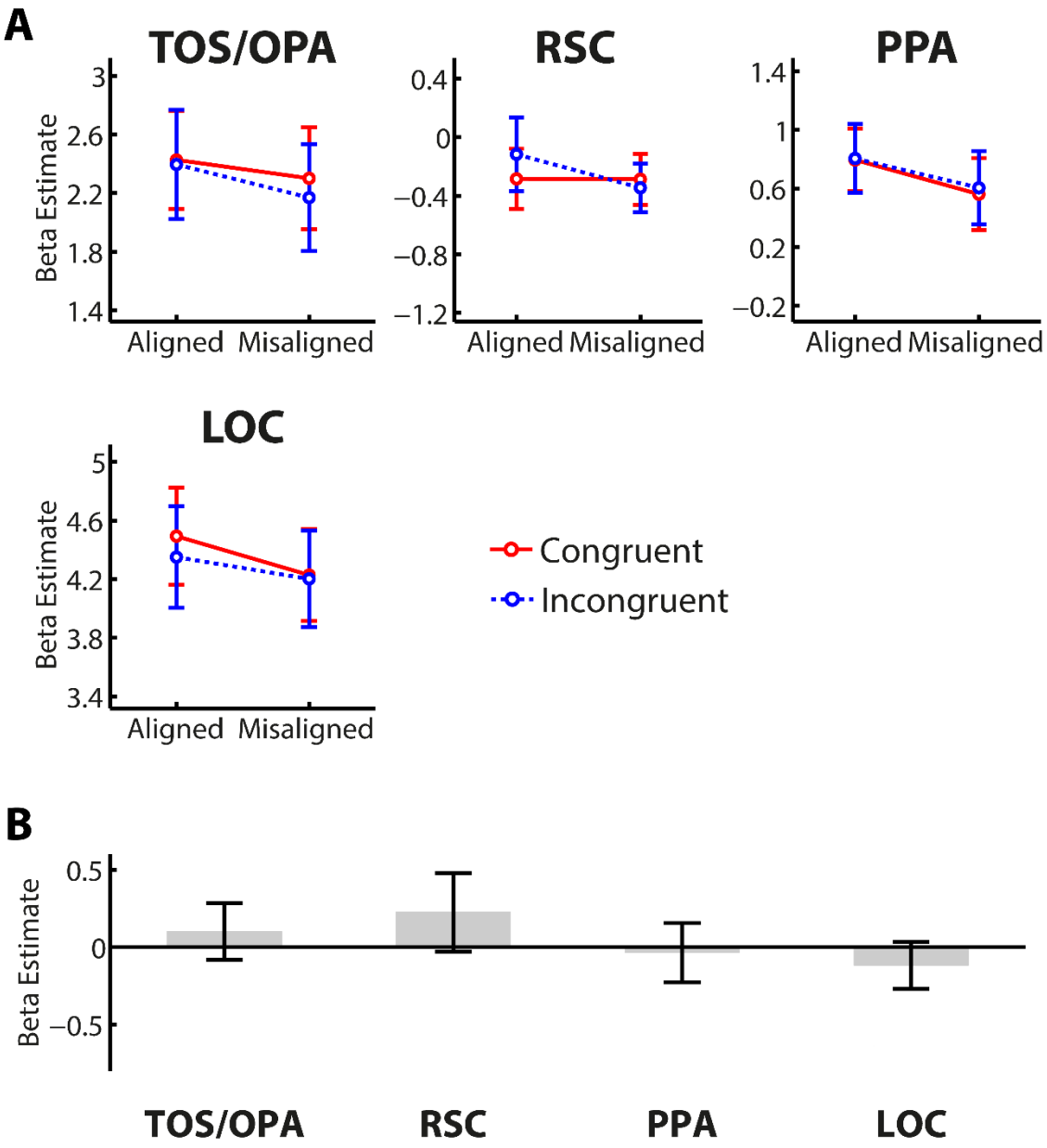

Figure 8. Neural responses to the top-different conditions of the composite task in object- and scene-responsive ROIs. (A) shows neural responses as a function of alignment and congruency. (B) shows the interaction effect between alignment and congruency. Error bars indicate ±1 SEM.

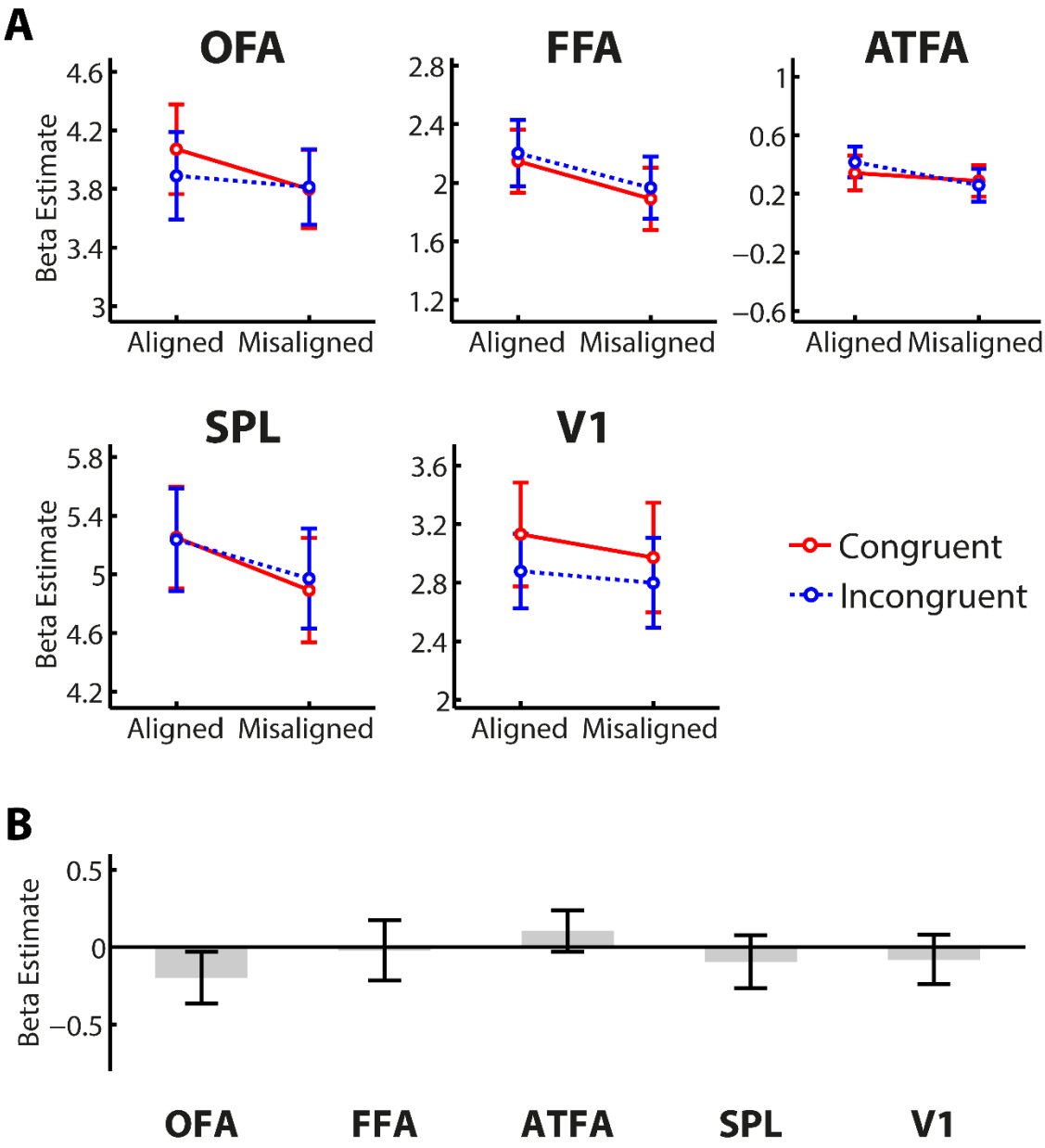

Figure 9. Neural responses to the top-different conditions of the composite task in face-responsive, perceptual-grouping related and early visual ROIs. (A) shows neural responses as a function of alignment and congruency. (B) shows the interaction effect between alignment and congruency. Error bars indicate ±1 SEM.

6.4. Discussion

In this study, we investigated the neural correlates of holistic processing of non-expertise line pattern objects. Despite our behavioural results showing strong evidence of holistic processing of the line pattern objects, we were unable to identify any regions showing significant neural responses related to this holistic processing. Furthermore, as previous studies investigating holistic processing of faces using a similar experimental design found neural responses related to holistic processing in several brain regions (Schiltz et al., 2010; our own work in Chapter 5), we hypothesize this null result could be due to differences between the neural processes involved in holistic processing of line pattern objects and faces. Such differences could involve a weaker neural signal that we could not detect using our experimental fMRI design or involvement of other brain regions outside of the areas we localised in this study.

Recent behavioural studies have also suggested that different mechanisms may underlie holistic processing of faces and line pattern objects. One study tested if there is interference between the processing of faces and line patterns when they are processed concurrently in an paradigm interleaving faces and line patterns (Curby, Huang, & Moerel, 2019). Using this design, the authors found no evidence of interference, suggesting that different mechanisms underlie holistic processing of faces and line patterns. However, another study that investigated holistic processing of overlaid faces and line pattern objects did find interference between the two stimuli, suggesting there may be similar mechanisms involved in processing the two types of stimuli at a perceptual level (Curby & Moerel, 2019). The authors concluded from these two studies that early, perceptual mechanisms relating to holistic processing may overlap between faces and line patterns, whereas aspects of holistic processing related to experience may be specific to faces and objects of expertise. This conclusion is further supported by differences in the effect of inversion between faces and line pattern objects. Holistic processing is strongly reduced by inversion for faces and expertise objects (Gauthier & Tarr, 1997; Tanaka & Farah, 1993; Young et al., 1987), whereas inversion does not affect the Gestalt connections of the line pattern objects. In Chapter 5, we found that the differences in neural activation between aligned and misaligned faces in the LOC and FFA mirror previous findings of neural response differences

between upright and inverted faces in these regions (Aguirre, Singh, & D'Esposito, 1999; Epstein, Higgins, Parker, Aguirre, & Cooperman, 2006; Grotheer & Kovacs, 2014; Haxby et al., 1999; Yovel & Kanwisher, 2005). In this study, we found no differences between neural responses to aligned and misaligned line patterns in these regions, suggesting the neural alignment differences in LOC and FFA may be related to expertise.

We were unable to identify any regions showing significant neural responses related to holistic processing of the line pattern objects in this study. Therefore, we cannot conclude whether regions outside of those we localized might be involved, or perhaps if the neural responses were present in the regions we localized but too weak to be detected with our experimental fMRI paradigm. We identified a non-significant trend for higher responses to aligned as compared to misaligned line pattern objects in the SPL, suggesting there could be involvement of parietal cortex in holistic processing of the line patterns. A larger sample size, or different localization method might be able to uncover an effect in the SPL. It is possible that other brain regions that we did not localize might be involved in holistic processing of the line patterns. For example, studies investigating neural responses to global Gestalt have found responses in several regions of parietal and parietal-occipital cortex, in particular at the location of the temporo-parietal junction (Huberle & Karnath, 2012; Rennig, Bilalić, Huberle, Karnath, & Himmelbach, 2013; Rennig, Himmelbach, Huberle, & Karnath, 2015; Seymour, Karnath, & Himmelbach, 2008). Furthermore damage to these regions has been associated with simultanagnosia, a disorder affecting the perception of global Gestalt (da Silva, Millington, Bridge, James-Galton, & Plant, 2017). Thus it is possible that further work, localizing more regions associated with Gestalt perception might uncover the neural correlates of holistic processing of line pattern objects. Furthermore, future work could include a larger sample size, and use a block design to increase signal to noise ratio (Soares et al., 2016).

To conclude, in this study we show that despite a strong behavioural holistic processing effect for non-expertise line pattern objects, we were unable to identify any brain regions showing neural activity consistent with this behavioural effect. Thus, it is possible that the brain regions involved in holistic processing of non-expertise objects are different to those involved in holistic processing of faces, consistent with recent behavioural findings. Future work investigating holistic processing of non-expertise objects may be able

to uncover the neural correlates by localizing more regions, using larger sample sizes or using block-design fMRI paradigms.

Acknowledgements

This research was supported by the Max Planck Society, Germany.

References

Aguirre, G. K. (2007). Continuous carry-over designs for fMRI. *NeuroImage*, *35*(4), 1480–1494. https://doi.org/10.1016/j.neuroimage.2007.02.005

Aguirre, G. K., Singh, R., & D'Esposito, M. (1999). Stimulus inversion and the responses of face and object-sensitive cortical areas. *NeuroReport*, *10*(1), 189–194. https://doi.org/10.1097/00001756-199901180-00036

Altmann, C. F., Bülthoff, H. H., & Kourtzi, Z. (2003). Perceptual organization of local elements into global shapes in the human visual cortex. *Current Biology*, *13*(4), 342–349. https://doi.org/10.1016/S0960-9822(03)00052-6

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. *Journal of Neuroscience*, *30*(9), 3544–3552. https://doi.org/10.1523/JNEUROSCI.4863-09.2010

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Brandman, T., & Yovel, G. (2016). Bodies are Represented as Wholes Rather Than Their Sum of Parts in the Occipital-Temporal Cortex. *Cerebral Cortex*, *26*(2), 530–543. https://doi.org/10.1093/cercor/bhu205

Brooks, J. L. (2012). Counterbalancing for serial order carryover effects in experimental condition orders. *Psychological Methods*, *17*(4), 600–614. https://doi.org/10.1037/a0029310

Bukach, C. M., Phillips, W. S., & Gauthier, I. (2010). Limits of generalization between categories and implications for theories of category specificity. *Attention, Perception, & Psychophysics*, *72*(7), 1865–1874. https://doi.org/10.3758/APP.72.7.1865

Curby, K. M., Huang, M., & Moerel, D. (2019). Multiple paths to holistic processing: Holistic processing of Gestalt stimuli do not overlap with holistic face processing in the same manner as do objects of expertise. *Attention, Perception, and Psychophysics*, *81*(3), 716–726. https://doi.org/10.3758/s13414-018-01643-x

Curby, K. M., & Moerel, D. (2019). Behind the face of holistic perception: Holistic processing of Gestalt stimuli and faces recruit overlapping perceptual mechanisms. *Attention, Perception, and Psychophysics*, *81*(8), 2873–2880. https://doi.org/10.3758/s13414-019-01749-w

da Silva, M. N. M., Millington, R. S., Bridge, H., James-Galton, M., & Plant, G. T. (2017). Visual dysfunction in posterior cortical atrophy. *Frontiers in Neurology*, *8*(AUG), 1–11. https://doi.org/10.3389/fneur.2017.00389

Diamond, R., & Carey, S. (1986). Why faces are and are not special: an effect of expertise. *Journal of Experimental Psychology: General*, *115*(2), 107–117. https://doi.org/10.1037//0096-3445.115.2.107

Epstein, R., Higgins, J. S., Parker, W., Aguirre, G. K., & Cooperman, S. (2006). Cortical correlates of face and scene inversion: A comparison. *Neuropsychologia*, *44*(7), 1145–1158. https://doi.org/10.1016/j.neuropsychologia.2005.10.009

Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, *392*(6676), 598–601. https://doi.org/10.1038/33402

Friston, K. J., Rotshtein, P., Geng, J. J., Sterzer, P., & Henson, R. N. (2006). A critique of functional localisers. *NeuroImage*, *30*(4), 1077–1087. https://doi.org/10.1016/j.neuroimage.2005.08.012

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience*, *3*(2), 191–197. https://doi.org/10.1038/72140

Gauthier, I., & Tarr, M. J. (1997). Becoming a "Greeble" expert: Exploring mechanisms for face recognition.

Vision Research, *37*(12), 1673–1682. https://doi.org/10.1016/S0042-6989(96)00286-6

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), 568–573.

Gauthier, I., Tarr, M. J., Moylan, J., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). The fusiform "face area" is part of a network that processes faces at the individual level. *Journal of Cognitive Neuroscience*, *12*(3), 495–504. https://doi.org/10.1162/089892900562165

Goffaux, V. (2012). The discriminability of local cues determines the strength of holistic face processing. *Vision Research*, *64*, 17–22. https://doi.org/10.1016/j.visres.2012.04.022

Goffaux, V., Schiltz, C., Mur, M., & Goebel, R. (2013). Local discriminability determines the strength of holistic processing for faces in the fusiform face area. *Frontiers in Psychology*, *3*(JAN), 1–14. https://doi.org/10.3389/fpsyg.2012.00604

Grassi, P. R., Zaretskaya, N., & Bartels, A. (2016). Parietal cortex mediates perceptual Gestalt grouping independent of stimulus size. *NeuroImage*, *133*, 367–377. https://doi.org/10.1016/j.neuroimage.2016.03.008

Grassi, P. R., Zaretskaya, N., & Bartels, A. (2018). A Generic Mechanism for Perceptual Organization in the Parietal Cortex. *Journal of Neuroscience*, *38*(32), 7158–7169. https://doi.org/10.1523/JNEUROSCI.0436-18.2018

Grill-Spector, K. (2003). The neural basis of object perception. *Current Opinion in Neurobiology*, *13*(2), 159–166. https://doi.org/10.1016/S0959-4388(03)00040-0

Grotheer, M., & Kovacs, G. (2014). Repetition Probability Effects Depend on Prior Experiences. *Journal of Neuroscience*, *34*(19), 6640–6646. https://doi.org/10.1523/jneurosci.5326-13.2014

Harris, A., & Aguirre, G. K. (2008). The representation of parts and wholes in face-selective cortex. *Journal of Cognitive Neuroscience*, *20*(5), 863–878. https://doi.org/10.1162/jocn.2008.20509

Harris, A., & Aguirre, G. K. (2010). Neural Tuning for Face Wholes and Parts in Human Fusiform Gyrus Revealed by fMRI Adaptation. *Journal of Neurophysiology*, *104*(1), 336–345. https://doi.org/10.1152/jn.00626.2009

Haxby, J. V., Ungerleider, L. G., Clark, V. P., Schouten, J. L., Hoffman, E. A., & Martin, A. (1999). The effect of face inversion on activity in human neural systems for face and object perception. *Neuron*, *22*(1), 189–199. https://doi.org/10.1016/S0896-6273(00)80690-X

Hinds, O., Polimeni, J. R., Rajendran, N., Balasubramanian, M., Amunts, K., Zilles, K., … Triantafyllou, C. (2009). Locating the functional and anatomical boundaries of human primary visual cortex. *NeuroImage*, *46*(4), 915–922. https://doi.org/10.1016/j.neuroimage.2009.03.036

Hole, G. J. (1994). Configurational factors in the perception of unfamiliar faces. *Perception*, *23*(1), 65–74. https://doi.org/10.1068/p230065

Huberle, E., & Karnath, H. O. (2012). The role of temporo-parietal junction (TPJ) in global Gestalt perception. *Brain Structure and Function*, *217*(3), 735–746. https://doi.org/10.1007/s00429-011-0369-y

Kamps, F. S., Julian, J. B., Kubilius, J., Kanwisher, N., & Dilks, D. D. (2016). The occipital place area represents the local elements of scenes. *NeuroImage*, *132*, 417–424. https://doi.org/10.1016/j.neuroimage.2016.02.062

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, *17*(11), 4302–4311.

Kleiner, M., Brainard, D., & Pelli, D. (2007). "What's new in Psychtoolbox-3?" In *Perception 36 ECVP Abstract Supplement.*

Liu, J., Harris, A., & Kanwisher, N. (2010). Perception of face parts and face configurations: an fMRI study. *Journal of Cognitive Neuroscience*, *22*(1), 203–211. https://doi.org/10.1162/jocn.2009.21203

Maguire, E. (2001). The retrosplenial contribution to human navigation: A review of lesion and neuroimaging findings. *Scandinavian Journal of Psychology*, *42*(3), 225–238. https://doi.org/10.1111/1467-9450.00233

Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., … Tootell, R. B. H. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(18), 8135–8139. https://doi.org/10.1073/pnas.92.18.8135

Rajimehr, R., Young, J. C., & Tootell, R. B. H. (2009). An anterior temporal face patch in human cortex, predicted by macaque maps. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(6), 1995–2000. https://doi.org/10.1073/pnas.0807304106

Rennig, J., Bilalić, M., Huberle, E., Karnath, H.-O., & Himmelbach, M. (2013). The temporo-parietal junction contributes to global gestalt perception-evidence from studies in chess experts. *Frontiers in Human Neuroscience*, *7*(513), 1–11. https://doi.org/10.3389/fnhum.2013.00513

Rennig, J., Himmelbach, M., Huberle, E., & Karnath, H.-O. (2015). Involvement of the TPJ Area in Processing of Novel Global Forms. *Journal of Cognitive Neuroscience*, *27*(8), 1587–1600. https://doi.org/10.1162/jocn_a_00809

Ross, D. A., Tamber-Rosenau, B. J., Palermi, T. J., Zhang, J., Xu, Y., & Gauthier, I. (2018). High-resolution Functional Magnetic Resonance Imaging Reveals Configural Processing of Cars in Right Anterior Fusiform Face Area of Car Experts. *Journal of Cognitive Neuroscience*, *30*(7), 973–984. https://doi.org/10.1162/jocn_a_01256

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), 1–16. https://doi.org/10.1167/10.2.25

Schiltz, C., & Rossion, B. (2006). Faces are represented holistically in the human occipito-temporal cortex. *NeuroImage*, *32*(3), 1385–1394. https://doi.org/10.1016/j.neuroimage.2006.05.037

Schindler, A., & Bartels, A. (2016). Visual high-level regions respond to high-level stimulus content in the absence of low-level confounds. *NeuroImage*, *132*, 520–525. https://doi.org/10.1016/j.neuroimage.2016.03.011

Seymour, K., Karnath, H.-O., & Himmelbach, M. (2008). Perceptual grouping in the human brain: Common processing of different cues. *NeuroReport*, *19*(18), 1769–1772. https://doi.org/10.1097/WNR.0b013e328318ed82

Soares, J. M., Magalhães, R., Moreira, P. S., Sousa, A., Ganz, E., Sampaio, A., … Sousa, N. (2016). A Hitchhiker's guide to functional magnetic resonance imaging. *Frontiers in Neuroscience*, *10*(515), 1–35. https://doi.org/10.3389/fnins.2016.00515

Tanaka, J. W., & Farah, M. J. (1993). Parts and Wholes in Face Recognition. *The Quarterly Journal of Experimental Psychology Section A*, *46*(2), 225–245. https://doi.org/10.1080/14640749308401045

Tsao, D. Y., Moeller, S., & Freiwald, W. A. (2008). Comparing face patch systems in macaques and humans. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(49), 19514–19519. https://doi.org/10.1073/pnas.0809662105

Wong, A. C.-N., Palmeri, T. J., & Gauthier, I. (2009). Conditions for facelike expertise with objects: Becoming a ziggerin expert - but which type? *Psychological Science*, *20*(9), 1108–1117. https://doi.org/10.1111/j.1467-9280.2009.02430.x

Xu, Y. (2005). Revisiting the role of the fusiform face area in visual expertise. *Cerebral Cortex*, *15*(8), 1234–1242. https://doi.org/10.1093/cercor/bhi006

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception*, *16*, 747–759. https://doi.org/10.1068/p160747n

Yovel, G., & Kanwisher, N. (2005). The neural basis of the behavioral face-inversion effect. *Current Biology*,

15(24), 2256–2262. https://doi.org/10.1016/j.cub.2005.10.072

Zaretskaya, N., Anstis, S., & Bartels, A. (2013). Parietal Cortex Mediates Conscious Perception of Illusory Gestalt. *The Journal of Neuroscience*, *33*(2), 523–531. https://doi.org/10.1523/JNEUROSCI.2905-12.2013

Zhao, M., Bülthoff, H. H., & Bülthoff, I. (2016). Beyond faces and expertise: Face-like holistic processing of nonface objects in the absence of expertise. *Psychological Science*, *27*(2), 213–222. https://doi.org/10.1177/0956797615617779

Zhao, M., Cheung, S.-H., Wong, A. C.-N., Rhodes, G., Chan, E. K. S., Chan, W. W. L., & Hayward, W. G. (2014). Processing of configural and componential information in face-selective cortical areas. *Cognitive Neuroscience*, *5*(3–4), 160–167. https://doi.org/10.1080/17588928.2014.912207

Acknowledgements

I would like to thank my supervisors Isabelle Bülthoff, Mintao Zhao and Andreas Bartels for their excellent advice and support throughout my doctoral work. I would like to thank Heinrich Bülthoff for providing me with the opportunity to conduct my doctorate at the Max Planck Institute for Biological Cybernetics. I greatly enjoyed the opportunity to work with Michael Black, Timo Bolkart and Javier Romero at the Max Planck Institute for Intelligent Systems, and I thank them for all the insights they provided me into body modelling. I would also like to thank the other members of my advisory board, Marc Himmelbach and Betty Mohler for their advice during my doctorate.

I would like to thank my colleagues at the Max Planck Institute for Biological Cybernetics and the Centre for Integrative Neuroscience, in particular those who assisted me as scan buddies during my studies, Junsuk Kim, Renée Hartig, Adamantini Chatzipanagioti, Xiaoning Zhao, Gizem Altan, Michael Bannert, Fabienne Schlüsener, Pablo Grassi, Oleksandra Shevtsova, Duangkamol Srismith, Didem Korkmaz Hacialihafiz, Soyoung Kwon, Georg Schauer and Andreas Schindler. I would also like to thank Rebekka Tenderra, Melissa Fuchs and Cristina Pelea for their hard work during their lab rotation and internship projects with me.

I would like to thank my friends and family for their support during my doctorate. Finally, and most importantly, I would like to thank the participants of my studies. Without their hard work none of these studies would have been possible.