52 The Functional Organization of the Ventral Visual Pathway in Humans

NANCY KANWISHER AND DANIEL D. DILKS

We can recognize an object within a fraction of a second, even if we have never seen that particular object before, and even if we have no advance information about what kind of object it might be (Potter, 1976; Thorpe, Fize, & Marlot, 1996). The cognitive and neural mechanisms underlying this remarkable ability are not well understood, and current computer vision algorithms still lag far behind human performance. One promising strategy for attempting to understand human visual recognition is to characterize the neural system that accomplishes it, the ventral visual pathway (VVP), which extends from the occipital lobe into inferior and lateral regions of the temporal lobe. Here we describe key results from the last 15 years of neuroimaging research on humans that have begun to elucidate the general organization and functional properties of the cortical regions involved in visually perceiving people, places, and things.

The central finding from this now-substantial body of work is that the VVP is not homogeneous but is instead a highly differentiated structure containing a set of regions each with its own distinct functional profile. These regions include the fusiform face area (FFA), which responds selectively to faces (Kanwisher, McDermott, & Chun, 1997a; McCarthy et al., 1997), the parahippocampal place area (PPA), which responds selectively to places (Epstein & Kanwisher, 1998), the extrastriate body area (EBA), which responds selectively to bodies (Downing et al., 2001), the lateral occipital complex (LOC), which responds to object shape (Kanwisher et al., 1997b; Malach et al., 1995) largely independent of object category, and the visual word form area (VWFA), which responds selectively to both visually presented words and consonant strings (Baker, Hutchison, & Kanwisher, 2007; Cohen et al., 2000). Each of these regions is present in approximately the same location in virtually every healthy subject. These regions and their cohorts (e.g., the occipital face area or OFA) constitute the fundamental machinery of high-level visual recognition in humans. An understanding of the function of each of these regions is likely to provide important clues about how visual recognition works.

In this chapter we first describe the best-established functionally specific regions in the VVP and then the ongoing controversies about the degree of such specialization. We then review the available evidence on the functional organization of the entire VVP, and finally, we consider the computational advantages that may be afforded by having specialized regions in the first place.

THE BEST-ESTABLISHED FUNCTIONALLY SPECIFIC REGIONS IN THE VENTRAL VISUAL PATHWAY

Figure 52.1 shows the main functionally distinct regions of the VVP. We begin by focusing on the FFA, PPA, EBA, and LOC because these are the best-established regions in this pathway in several respects. First, each of these regions has been found consistently in a large number of studies and labs; although the theoretical significance of these regions can be debated, their existence cannot. Second, the category selectivity by which each region is defined is not merely statistically significant but also large in effect size: Each of these regions responds about twice as strongly to any stimulus from its preferred category as to any nonpreferred stimulus category. Effect size is often ignored in the brain imaging literature, but it should not be, as it determines the strength of the inference you can draw: If you know how to double the response of a region, you generally have a better handle on its function than if you merely know how to change its response by a small amount. Third, the fact that these regions can be found easily and now algorithmically (Julian et al., 2012) in any healthy subject has made possible a "region of interest" (ROI) research strategy whereby the region is first functionally identified in each subject individually in a short "localizer" scan and then characterized in rich detail in subsequent experiments, including not just its response profile but the information it represents.

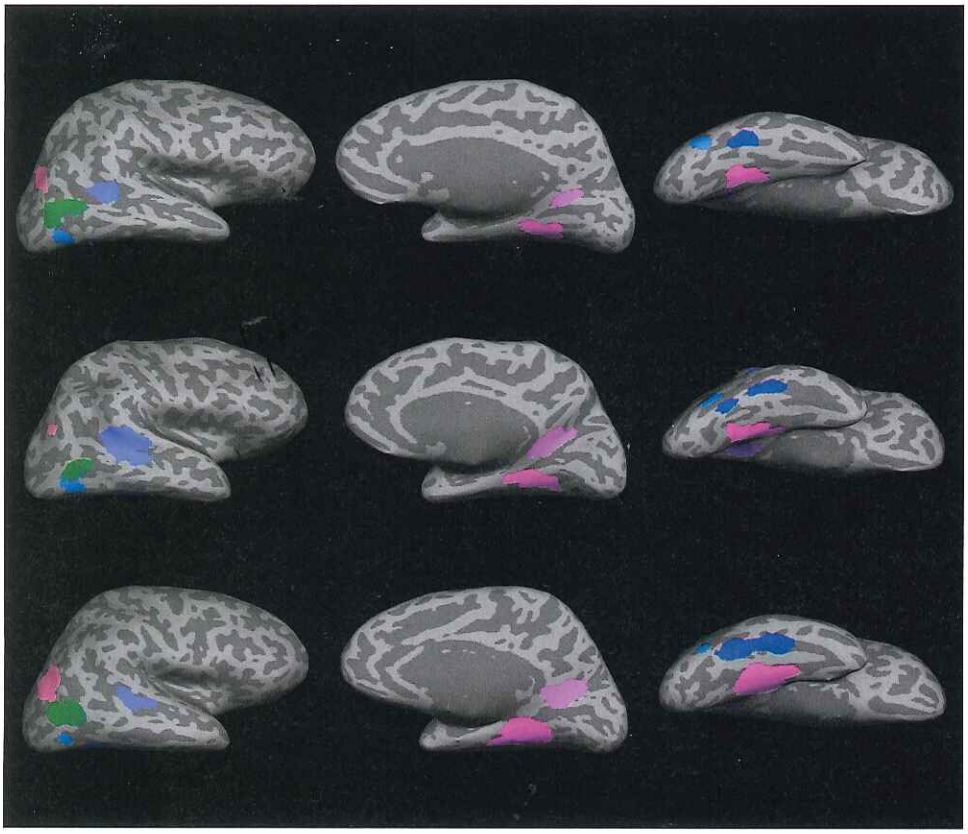


FIGURE 52.1 Inflated brain from three representative subjects showing regions specifically involved in the perception of faces (blues), places (pinks), and bodies (green). Dark blue, FFA; purplish blue, pSTS; light blue, OFA; magenta, PPA; light purple/pink, RSC; reddish pink, OPA; green, EBA. Each of these regions can be found in a short functional scan in essentially every healthy subject. LOC is not shown here; it is a very large region generally responsive to any object shape, and hence both of its subregions (LO and pFs) overlap partially with some of the regions shown here. The VWFA (left hemisphere) and FBA (partially overlapping with FFA) are also not pictured.

The approach to neuroimaging in which cognitive functions are assigned to brain regions has been widely maligned as "mere phrenology," as if the label itself is an argument against the whole enterprise. But name-calling is not argumentation any more in science than in the schoolyard. There is nothing fundamentally wrongheaded with the effort to characterize the cognitive functions of particular brain regions—it is an empirical question how successful this research program will turn out to be. The problem is rather that most neuroimaging studies neither robustly establish the existence of a functionally specific region nor precisely identify its function. When a brain region is identified based on just a few conditions in a single study, the evidence will necessarily be weak, and the functional characterization of the region will be sparse and unsatisfying. But when the same brain region has been identified in many labs, and when each of those labs has measured the functional response of that region in numerous conditions, each testing a different hypothesis about the representations it contains, then we can begin to approach the kind of rich cognitive (and perhaps some day computational) characterization of that brain region that will be of real theoretical significance to cognitive science. Thus, brain imaging can discover not just the *locations* of cognitive functions (who cares?) but, more fundamentally, it can discover and characterize those cognitive functions themselves. That is, brain imaging can exploit the functional specificity of the brain to discover the cognitive architecture of the mind. Although still very much a work in progress, this enterprise is most advanced for the regions we begin with here: the FFA, PPA, EBA, and LOC.

The Fusiform Face Area

The FFA is the region of the midfusiform gyrus (on the bottom surface of the cerebral cortex, just above the cerebellum) that responds significantly more strongly when subjects view faces than when they view objects

(Kanwisher, McDermott, & Chun, 1997a; McCarthy et al., 1997; Puce et al., 1996). The precise shape of the FFA varies across subjects, and consists of two regions in some subjects (Weiner & Grill-Spector, 2012), but is present in nearly all subjects and highly reproducible within each (Peelen & Downing, 2005). Earlier neuroimaging work (Haxby et al., 1994; Puce et al., 1995) had shown activations in this general region for faces (compared to scrambled faces, textures, and letter strings), but those contrasts left open the possibility that this region was engaged in processing generic object shape. The subsequent evidence for the *specificity* of this region for faces per se came from the discovery that this region responds similarly and strongly to a wide variety of face images (Kanwisher & Yovel, 2006), including photos of familiar and unfamiliar faces, schematic faces, cartoon faces, and cat faces (Kanwisher & Barton, 2010) as well as faces presented in different sizes, locations, and viewpoints (Axelrod & Yovel, 2012; Grill-Spector et al., 1999; Schwarzlose et al., 2008) and much less strongly to nonface stimuli. Extensive evidence now rejects alternative hypotheses proposed earlier that the FFA is more generally engaged in fine-grained discrimination of exemplars of any category, or of any category for which the subject has gained substantial expertise (McKone & Robbins, 2010; Tarr & Gauthier, 2000; Yovel & Kanwisher, 2004). Consistent with the evidence from fMRI, face-selective responses have been observed in approximately the same location in subdural electrode recordings from the brains of subjects undergoing presurgical mapping for epilepsy treatment (Allison et al., 1999; McCarthy et al., 1999; Puce, Allison, & McCarthy, 1999), and lesions in approximately this location can produce selective deficits in face perception (Kanwisher & Barton, 2010). Thus, as discussed further below, the FFA appears to be quite selectively engaged during the perception of faces.

Answering the question of what exactly the FFA does with faces has been more difficult. Importantly, the magnitude of the FFA response is correlated trial-by-trial with success both in detecting the presence of faces and in identifying individual faces (Grill-Spector & Malach, 2004) but not with detection or identification of nonface objects, and the extent of face selectivity in the fusiform is correlated with behavioral ability at face (but not object) identification across subjects (Furl et al., 2011). In terms of the kind of information represented in the FFA, current evidence indicates that the FFA is sensitive to multiple aspects of face stimuli including face parts (eyes, noses, and mouths), the T-shaped configuration of those features, and external features of faces (Liu, Harris, & Kanwisher, 2010). Further, FFA responses show some invariance across changes in stimulus position and image size (Schwarzlose et al., 2008; but see Yue et al., 2011). Using multivoxel pattern analysis (MVPA), a recent study (Axelrod & Yovel, 2012) found similar neural representations for mirror-symmetric views of faces but not for other changes in viewpoint, and other studies using fMRI adaptation have found a moderate invariance for viewpoint changes (up to 30°) for familiar (celebrity) faces (e.g., Ewbank & Andrews, 2008). Finally, the FFA exhibits neural correlates of well-established behavioral signatures of face perception (McKone & Robbins, 2010), including sensitivity to differences in face identity for upright but not inverted faces (Yovel & Kanwisher, 2005) and sensitivity to holistic information in upright but not inverted faces (Schiltz & Rossion, 2006). Thus, the FFA appears to represent perceptual information about face shape in a fashion partially invariant to image changes and to reflect the well-known behavioral signatures of face-specific processing.

The Parahippocampal Place Area

The PPA is defined functionally as the region adjacent to the collateral sulcus in parahippocampal cortex that responds significantly more strongly to images of scenes than objects (Epstein & Kanwisher, 1998). The PPA responds to a wide variety of scenes, including indoor and outdoor scenes, familiar and unfamiliar scenes, and even abstract "scenes" made of Legos (Aguirre et al., 1996; Epstein, 2005, 2008). The PPA is primarily responsive to the spatial layout of one's surroundings: Its response is not reduced when all of the objects are removed from an indoor scene, leaving just the floor and walls (Epstein & Kanwisher, 1998). The PPA has also been shown to respond selectively to high-spatial-frequency geometric shapes in humans and monkeys (Rajimehr et al., 2011), which suggests that the PPA may use such information for detecting scene details during place perception and navigation. By contrast, Bar and colleagues have proposed that the parahippocampal response to scenes does not reflect spatial layout but rather the activation of a "context frame" representation that includes information about which objects typically appear in that context and where they are likely to be located relative to each other (Bar, 2004). However, the evidence for the context hypothesis is weak: The finding of greater response to strong- versus weak-context objects only replicates at slow presentation rates, is reliable in only a minority of subjects, and can be alternatively explained in terms of scene imagery (Epstein & Ward, 2010). On the other hand, consistent with the spatial layout hypothesis, studies using MVPA found that the PPA contains significantly more

information about spatial layout—that is, whether a scene is "open" versus "closed" (Oliva & Torralba, 2001)—than information about whether a scene is manmade (e.g., a city) or natural (e.g., a forest) (Kravitz, Peng, & Baker, 2011; Park et al., 2011). Evidence that the PPA is not only activated when information about spatial layout is processed but that it is further *necessary* for this function comes from patients with damage in or near the PPA, who have deficits in simple identification of scenes or landmarks (Aguirre & D'Esposito, 1999; Mendez & Cherrier, 2003) and difficulty more generally in knowing where they are (Epstein et al., 2001; Habib & Sirigu, 1987). This high response to spatial layout information is tantalizingly reminiscent of the "geometric module" (Cheng & Gallistel, 1984; Hermer & Spelke, 1996), inferred from behavioral data, in which rats and human infants (and adults whose language system is tied up by a concurrent verbal task) rely exclusively on the layout of space, not on objects or landmarks, to reorient themselves in an environment once they are disoriented. However, we recently found that representations in the PPA are largely invariant to mirror-image reversals, a result that challenges its role in navigation and reorientation (Dilks et al., 2011). The precise role of the PPA in place perception and navigation is a topic of ongoing investigation (see chapter 51 by Oliva).

The Extrastriate Body Area

The EBA is a region on the lateral surface of the brain adjacent to (and sometimes partly overlapping with) visual motion area MT, which responds significantly more strongly to images of bodies than to images of objects or faces (Downing et al., 2001). This region responds equally to visually different images of bodies, from a photograph of a hand, to a photograph of a body (human or animal), to a line drawing, or even a schematic stick figure of a person. The EBA is more involved in perceiving other people's bodies, in a viewpoint-dependent manner (Taylor, Wiggett, & Downing, 2010), than one's own (Chan, Peelen, & Downing, 2004; Saxe, Jamal, & Powell, 2006) and is more engaged in the perception of the form/identity of bodies than in the actions they are carrying out (Downing & Peelen, 2011; Moro et al., 2008). Evidence that this region not only is activated during, but is also necessary for, the perception of bodies comes from studies in which disruption of the EBA by a brain lesion (Moro et al., 2008) or by transcranial magnetic stimulation (TMS) (Pitcher et al., 2009; Urgesi, Berlucchi, & Aglioti, 2004; van Koningsbruggen, Peelen, & Downing, 2013) impairs the perception of body form but not the perception of faces or of object shape (Pitcher et al., 2009). Recent evidence indicates that the EBA may consist of several subregions and that these subregions of the EBA and neighboring cortex may respond differentially to distinct body parts (Bracci et al., 2010; Op de Beeck et al., 2010; Orlov, Makin, & Zohary, 2010; Weiner & Grill-Spector, 2011). A recent review (Downing & Peelen, 2011) argues that the EBA and another more ventral body-specific region, the fusiform body area or FBA (Schwarzlose, Baker, & Kanwisher, 2005), are not primarily engaged in a higher-level interpretation of the individual identity, emotional content, motion, or action goals but rather extract a "cognitively unelaborated" visual representation of the shape and posture of the people in the current percept.

The Lateral Occipital Complex

In addition to the category-specific regions just described, a large region of lateral and inferior occipital cortex just anterior to retinotopic cortex (the lateral occipital complex or LOC), and partially overlapping with some of the regions described above, responds more strongly to stimuli depicting shapes than stimuli with similar low-level features that do not depict shapes (Kanwisher et al., 1997b; Malach et al., 1995). Common areas within LOC are activated by shapes defined by motion, texture, and luminance contours (Grill-Spector et al., 1998), showing that representations of shape in LOC are quite abstract. Importantly, the LOC responds similarly to familiar and unfamiliar shapes (Kanwisher et al., 1997b; Malach et al., 1995), suggesting that this region is not involved in matching to stored object representations or to semantic or verbal coding of the stimuli. Further support for the idea that LOC represents object shape, not semantic information about objects, comes from the fact that fMRI adaptation is not found in LOC across objects that are similar in meaning but differ in shape (Kim et al., 2009). Several studies have implicated the LOC in visual object recognition by showing that activity in this region is correlated with success on a variety of object recognition tasks (Bar et al., 2001; Grill-Spector et al., 2000), and indeed the location of LOC matches very nicely the location of the lesion in the famous ventral pathway agnosic patient DF (James et al., 2003). Thus, an investigation of the response properties of the LOC may provide important clues about the nature of the representations underlying object recognition.

MVPA and fMRI adaptation studies show that LOC contains fairly abstract representations of object shape. First, fMRI adaptation studies have shown that representations in the vicinity of LOC are partially invariant

to changes in size and position but largely specific to viewpoint and direction of illumination (Grill-Spector et al., 1999). Evidence on the nature of the shape representations in LOC comes from the fact that fMRI adaptation occurs in LOC between stimulus pairs that have different contours but the same perceived shape (because of changes in occlusion) but not between stimulus pairs with identical contours but different perceived shape (because of a figure–ground reversal) (Kourtzi & Kanwisher, 2001). Further, MVPA analyses show that a posterior subregion of LOC (often called LO) contains representations that are more tied to the stimulus, whereas a more anterior subregion (called pFs) contains representations that are correlated with observer-specific perceptions of shape similarity (Haushofer, Livingstone, & Kanwisher, 2008). Finally, another recent study found that the more posterior region LO showed sensitivity to mirror-image reversals of objects, although the more anterior region pFs did not, suggesting a hierarchy of object processing in which left–right information is represented at earlier (more posterior) stages in the hierarchy and invariance is then computed at later (more anterior) stages (Dilks et al., 2011).

In sum, the VVP contains a large multipart region, the LOC, that responds strongly to object structure but that exhibits little selectivity for specific object categories, along with a small number of category-specific regions (for faces, places, and bodies). Efforts to date have not found regions of the VVP robustly selective for other categories (Downing et al., 2006; Lashkari et al., 2011), with the exception of the VWFA (Baker et al., 2007; Cohen et al., 2000).

Specificity: Do Category-Selective Regions Contribute Only to the Perception of Their Preferred Stimuli?

So far, we have argued that the FFA, PPA, and EBA are primarily if not exclusively engaged in processing their "preferred" stimuli (the stimulus class they respond most strongly to). However, each of these regions responds significantly (albeit weakly) to objects that are not in the preferred category ("nonpreferred" objects). Further, in what we consider the most important challenge to the claimed specificity of these regions, Haxby and colleagues (Haxby et al., 2001) reported that the spatial *pattern* of response across the FFA contains information about nonfaces and that the pattern of response within the PPA contains information about nonscenes, and hence, that "Regions such as the 'PPA' and 'FFA' are not dedicated to representing only spatial arrangements or human faces, but, rather, are part of a more extended representation for all objects."

We too find that the FFA and PPA contain information about nonpreferred stimuli (Reddy & Kanwisher, 2007), and current physiological evidence (Tsao et al., 2006) also indicates that face-selective regions carry a small but significant amount of information about nonpreferred stimuli. However, the information that category-selective regions contain is much weaker for nonpreferred than preferred stimuli. As Haxby and colleagues (O'Toole et al., 2005) noted, "preferred regions for faces and houses are not well suited to object classifications that do not involve faces and houses, respectively." These findings raise two questions, which we address next.

First, is the information about nonpreferred objects in category-selective regions still present when subjects view cluttered displays more typical of real-world vision? Most analyses of the spatial pattern of the fMRI response within the VVP have been based on responses elicited by single cut-out objects on a blank background, presented at the fovea. Of course real-world visual stimuli are not this simple: A typical visual scene contains multiple objects and complex background textures (i.e., "clutter"). A recent study (Reddy & Kanwisher, 2007) tested whether information about nonpreferred objects is present in a minimalist case of clutter, with two objects present simultaneously in the visual field (both on blank backgrounds). When single cut-out stimuli were shown one at a time, the pattern of response in the FFA contained considerable information about faces and significant though weaker information about nonfaces. Similarly, the pattern of response in the PPA contained robust information about houses (which activate the PPA strongly, although not as strongly as a full scene) and significant but weak information about nonscenes. Crucially, however, when two objects were present at once, information about preferred stimuli was virtually undiminished from the single-object case, but information about nonpreferred stimuli dropped to insignificance (Reddy & Kanwisher, 2007). This study and later related studies suggest that category-selective regions may have little or no information about nonpreferred stimuli under natural (i.e., cluttered) viewing conditions.

Still, given that fMRI is bound to underestimate the information present in the full neural population code, it is possible that future physiological studies will reveal some information about nonpreferred stimuli in the FFA, the PPA, and similarly selective regions, even for the cluttered stimuli typical of real-world viewing. Thus, the second and most important question is whether such information is *used in the perception of those stimuli*, or whether it is epiphenomenal (Williams, Berberovic, & Mattingley, 2007). Some relevant evidence is available

for the case of the FFA from the study of individuals with focal brain damage. Some of these individuals exhibit deficits only in face perception (i.e., prosopagnosia), with little or no deficit in object recognition, after damage to regions in or near the FFA, suggesting that even if the FFA contains information about nonfaces, this information is not necessary for object perception. Although no published case of acquired prosopagnosia has completely ruled out the existence of any other deficits beyond face perception (Garrido, Duchaine, & Nakayama, 2008) using the most sensitive tests of object perception such as reaction time measures (Gauthier, Behrmann, & Tarr, 1999), some cases come close (Sergent & Signoret, 1992; Wada & Yamamoto, 2001).

However, because the locus and extent of lesions in humans are not under our control, an important complementary method for testing the functional specificity and causal role of cortical regions in perception is TMS. In TMS a brief magnetic pulse is delivered to the scalp through a coil held next to the scalp, disrupting neural processing in the cortical region immediately beneath the coil. We can now precisely position the TMS coil to directly target specific cortical regions defined functionally within individual subjects. Although the FFA and PPA are too medial to be reached by TMS, other more lateral category-selective regions (e.g., the face-selective OFA and the scene-selective OPA, previously known as TOS—Dilks et al., 2013—discussed in more detail below) can be. Using this method, Pitcher and colleagues (Pitcher et al., 2009) showed that TMS to the EBA disrupted perception of bodies (Urgesi, Berlucchi, & Aglioti, 2004) *but not faces or objects*, TMS to the OFA disrupted perception of faces *but not objects or bodies* (but see Silvanto et al., 2010), and TMS to LO disrupted perception of objects *but not bodies or faces*. Similarly, Dilks and colleagues (Dilks et al., 2013) recently showed that TMS to OPA impaired discrimination of scenes *but not faces*, while TMS to OFA impaired discrimination of faces *not scenes*. In a second experiment, Dilks et al. (2013) showed that TMS over OPA impaired categorization accuracy of scenes *but not objects*, while TMS over LOC impaired categorization accuracy of objects *not scenes* (see also Ganaden, Mullin, & Steeves, in press; Mullin & Steeves, 2011). These striking triple and double dissociations, respectively, suggest that category-selective regions play a causal role in the perception of their preferred stimulus class, but not their nonpreferred stimulus class. Thus, even if the pattern of response across these regions contains some information about nonpreferred stimulus categories, the available evidence suggests that such information plays no detectable causal role in perception.

In sum, current evidence suggests that category-selective regions sometimes contain weak but significant information about nonpreferred stimuli, which is likely to be underestimated by fMRI. Nonetheless, results from neuropsychology and TMS are consistent with the hypothesis that any information about nonpreferred stimuli in category-selective regions is epiphenomenal (i.e., not causally involved in perception of those stimuli). It will be important in the future to test this hypothesis further with new data from patients, TMS, and other disruption methods, such as electrical microstimulation in macaque monkeys and humans (Afraz, Kiani, & Esteky, 2006; Puce, Allison, & McCarthy, 1999).

THE FUNCTION AND STRUCTURE OF THE WHOLE VENTRAL VISUAL PATHWAY

Of course no complex cognitive process is accomplished in a single brain region, and arguments for the specificity of the regions described above in no way preclude an important role for other brain regions. "Earlier" cortical regions such as primary visual cortex are obviously crucial in the perception of faces, places, and bodies, and "higher" areas (e.g., in parietal and frontal regions) are also probably necessary for information in the FFA, PPA, and EBA to be used by other cognitive systems and to reach awareness (Kanwisher, 2001). Further, none of these regions is the only one with its defining selectivity. Other category-selective regions have not been studied in the same detail as the FFA, PPA, and EBA, so their functions are less clear. However, the existence of multiple selective regions, and the growing evidence for a functional division of labor between them, raises the exciting possibility that we may ultimately understand how face recognition, for example, emerges from the joint activity of a number of functionally distinct regions. Next, we briefly review the literature on other functionally distinct regions engaged in face and scene perception.

For faces, selective responses are found not only in the FFA but in many subjects also in a nearby but more posterior OFA (Gauthier et al., 2000) as well as other more anterior regions, such as the posterior superior temporal sulcus (pSTS) (Puce et al., 1998), and anterior temporal pole (Rajimehr, Young, & Tootell, 2009). Based on the more posterior location and generally lower selectivity of the OFA, it is often assumed to constitute an early stage of face perception (or face detection), which is then followed by continued processing in more anterior regions (e.g., the FFA and pSTS) (Haxby, Hoffman, & Gobbini, 2000; but see Rossion, Hanseeuw, & Dricot, 2012). Consistent with this picture, (1) the right OFA has been more implicated in the

representation of the parts of a face, including the eyes, nose, and mouth, than the configuration of those parts (Liu, Harris, & Kanwisher, 2010; Pitcher et al., 2007), whereas the FFA is sensitive to both face parts and their overall configuration in the face (Liu, Harris, & Kanwisher, 2010), and (2) the OFA is more sensitive to mirror-image reversals than is the FFA or pSTS (Axelrod & Yovel, 2012). By contrast, the face-selective region in the pSTS has been implicated in the representation of more dynamic high-level face and social information, including eye, mouth, and head movements (Carlin et al., 2011; Fox et al., 2009; Haxby, Hoffman, & Gobbini, 2000; Pitcher et al., 2011; Puce et al., 1998) and facial expression (Phillips & David, 1997; Winston et al., 2004). In one of the most striking functional dissociations within the face system, the face-selective region in the pSTS responds about three times as strongly to movies of faces (but not movies of bodies or objects) as to static snapshots taken from those face movies, whereas the FFA responds the same to movies and snapshots (Pitcher et al., 2011; see also Puce et al., 1998).

For scenes, selective responses are found not only in the PPA but also in retrosplenial complex (RSC) and a region near the transverse occipital sulcus, formerly referred to as "TOS" (Grill-Spector, 2003), now called the occipital place area (OPA) (Dilks et al., 2013). Like PPA, RSC is primarily responsive to the spatial layout of one's surroundings (Dilks et al., 2011; Epstein, 2008; Kravitz, Peng, & Baker, 2011; Park & Chun, 2009), with a recent study reporting only spatial layout information, not object information, in RSC (Harel, Kravitz, & Baker, 2013). However, in contrast to the ongoing debate about the PPA's precise role in place perception and navigation, most studies find clear evidence that RSC plays a role in navigation. For example, one fMRI adaptation study (Baumann & Mattingley, 2010) has shown that this region encodes heading direction. Further, RSC is sensitive to left–right information in scenes (i.e., mirror-image reversals of a scene), which is presumably important for navigation (Dilks et al., 2011). Finally, patients with RSC damage have been reported to recognize salient landmarks but not to use these landmarks to orient themselves or to navigate through a larger environment (Takahashi et al., 1997). OPA is the least-studied scene-selective region, but a recent study from our lab found that it is causally and selectively involved in scene perception: TMS over OPA impaired discrimination of scenes but not faces or objects (Dilks et al., 2013).

The existence of multiple face-selective regions and multiple scene-selective regions offers the exciting prospect of taking apart the process of face and scene perception by understanding the functional division of labor among the various regions within each system. An important part of this story concerns the connectivity of the different regions within each system and between those regions and the rest of the brain. Evidence on this important question is sparse, however, because neither of the two methods currently available in humans can answer these questions definitively. Resting-state fMRI is intriguing but can reveal strong correlations between regions known not to be directly connected (Tian et al., 2007). Diffusion tractography is subject to ambiguities both in tracing connectivity from specific functionally identified gray matter regions into the underlying white matter and in tracing specific connections through white matter (rather than simply following known major fiber bundles that run through the VVP, e.g., the ILF and IFOF). Indeed, preliminary evidence from these methods does not fully agree. Although both diffusion tractography and resting-state fMRI agree that the nearby OFA and FFA are connected, connections between the FFA and the face-selective region in the pSTS have been found with resting-state fMRI (Turk-Browne, Norman-Haignere, & McCarthy, 2010) but not with diffusion methods (Gschwind et al., 2011). It remains to be determined whether methods for discovering specific anatomical connections in humans will ultimately be able to discover the precise connectivity of specific subregions of the VVP.

Perhaps the biggest open question concerning the functional organization of the VVP is whether the functionally distinctive regions identified here are best thought of as discrete processors or whether it makes more sense to consider the broader region that contains them as a single processor in which each of these regions simply constitutes a local peak in the functional response. On the latter view, the question would still remain as to why that landscape would contain the particular replicable configuration it does across the VVP, and what if any are the dimensions represented by axes of this broader "map" (Op de Beeck, Haushofer, & Kanwisher, 2008; Kanwisher & Schwarzlose, 2008). Some of the locations of particular regions in the VVP may be explained in terms of a center–periphery map (Hasson et al., 2002) or a representation of real-world object size (Konkle, 2011). Further widely noted and intriguing aspects of the structure of the VVP are that many of the object- and category-selective regions come in pairs, with one on the ventral surface and one on the lateral surface (Hasson et al., 2003; Schwarzlose et al., 2008), and that body-selective and face-selective regions tend to be close and sometimes overlapping with each other (Kanwisher & Schwarzlose, 2008; Weiner & Grill-Spector, 2011). One clue to the question of whether the ventral pathway is best thought of as a

single representational space or a set of at least partially distinct processors may come from anatomy: To the extent that the different regions discussed above have distinctive connectivity and cytoarchitecture, that would support the interpretation of these regions as distinct entities. Indeed, recent evidence indicates that face selectivity within the fusiform gyrus can be predicted from connectivity with the rest of the brain (Saygin et al., 2011) and that at least one face-selective region in the fusiform gyrus may have a distinctive cytoarchitecture (Caspers et al., 2013).

WHY HAVE SELECTIVE REGIONS IN THE FIRST PLACE?

Perhaps the deepest question raised by the work on the VVP is this: Why do some visual categories get their own private piece of real estate in the VVP, whereas others apparently do not (Downing et al., 2006; Lashkari et al., 2011)? To think clearly about this question, we need to consider what computational advantages are afforded by functional specialization in the first place. To be detected by fMRI, functional specializations must have two properties: (1) selectivity of the response of neurons to the relevant information (e.g., face selectivity) and (2) spatial clustering of selective neurons. These phenomena are related but distinct (Ohki et al., 2005) and will be discussed in turn in this final, highly speculative section.

Selectivity/Sparseness

The advantages of selectivity, or "sparseness," in neural coding have been widely noted (Barlow, 1995; Foldiak & Young, 1995; Olshausen & Field, 2004). If a given object is coded by the activity of a small subset of the available neurons, then interference is minimized in two important senses. First, it is possible to represent multiple objects simultaneously with minimal ambiguity because the neural codes for different objects are unlikely to overlap. Thus, we can perceive a face and a place simultaneously without the two representations colliding (Reddy & Kanwisher, 2007). Perhaps this is one reason we have neural populations selectively responsive to faces, places, and bodies: to provide "private lines" of communication about particularly important classes of stimuli that are protected from crosstalk of other irrelevant information.

Second, the use of sparse codes can also reduce interference from learning, enabling us to learn new exemplars of one class of objects without altering stored information about another class of objects. With one neural population to represent faces and a nonoverlapping neural population to represent the spatial layout of places, we can learn new faces without disrupting our memories of places and vice versa. From this perspective we may expect to find relatively sparse codes for classes of information characterized by continual lifelong learning (such as faces and places).

Third, building specialized brain regions and precise connectivity linking them to other brain regions could bootstrap development by essentially hardwiring constraints on inductive inference. For example, if information in faces provides the key input required for learning about other people's minds, then perhaps the most efficient way to construct the machinery for thinking about other minds is to hardwire a face area and connect it to another available region, which will then have the constrained input it needs to construct the circuits necessary for social cognition. Evidence against this particular hypothesis comes from the recent finding that congenitally blind individuals show the same location and pattern of activation as sighted subjects when thinking about other people's thoughts, even though input from the FFA is likely very different or nonexistent in these people (Bedny, Pascual-Leone, & Saxe, 2009). Nevertheless, the general idea that specialized brain regions and their connections may serve as constraints on development is worth considering in other cases.

A fourth possible advantage of relatively sparse codes is metabolic rather than computational: Less energy is required if fewer neurons are firing. From this perspective the greatest lifelong energy savings would come about if sparse codes were available for classes of stimuli that occur most frequently (Foldiak & Young, 1995). Thus, even from a purely metabolic perspective, it makes sense to use relatively sparse codes for faces, places, and bodies because they are among the most frequently encountered visual stimuli.

In sum, sparse codes, in which information is represented by a relatively small percentage of the available neurons, each with relatively high selectivity, have certain advantages. At the same time, sparse codes have well-known disadvantages, such as greater susceptibility to damage (because of the smaller number of neurons involved in any given representation) and a smaller number of possible patterns that can be held (one at a time) by a fixed number of neurons. The speculation here is that these disadvantages are outweighed by the particular advantages in the coding of biologically important stimuli such as faces, places, and bodies: (1) reduction of interference or crosstalk when multiple stimuli must be represented simultaneously, (2) the ability to learn new information about one stimulus class without disrupting stored information about another class, (3) bootstrapping the development of

other regions, and (4) the potential energy efficiency of coding the most frequently encountered stimuli through the activity of the smallest number of neurons.

Spatial Clustering

The second property implied by functionally selective regions detected by fMRI, after selectivity of neurons, is spatial clustering of those neurons. Spatial clustering of functional properties is a familiar phenomenon in the brain, found not only in retinotopic, somatotopic, tonotopic, and other-topic maps that follow the organization of the receptor surface, but also in the organization of functional information that is computed de novo, such as orientation columns in primary visual cortex and chromotopic maps in posterior inferotemporal cortex (Conway & Tsao, 2009). Spatial organization is such a pervasive and familiar property of the cortex that we can easily forget to ask ourselves why it occurs. This mystery has been articulated most clearly (Chklovskii & Koulakov, 2004) as follows: "Imagine taking a cortical area containing a map and scrambling neurons in that area while preserving all the connections between neurons. Because the circuit remains unchanged, the functional properties of the neurons remain intact. Then the scrambled region without a map is functionally identical to the original one with the map." Given that the identical circuit can be constructed in a spatially clustered or spatially scrambled version, why does spatial clustering occur?

This question is sharpened by the facts that the strong spatial clustering seen in some systems, such as orientation-selective cells in cat visual cortex, is not found in other very similar systems, such as orientation-selective cells in rodent visual cortex (Ohki & Reid, 2007), and further by the fact that in the rodent olfactory processing pathway, the precise spatial clustering (and odorant specificity) constructed in the olfactory bulb is thrown away in the next stage of processing, the piriform cortex (Stettler & Axel, 2009).

Chklovskii and Koulakov (2004) argue that the need to minimize wiring length (for developmental, metabolic, and conduction delay reasons) must be a fundamental constraint in the nervous system that produces spatial clustering of neurons that are densely connected to each other. To the extent that this wiring-length minimization principle is an important determinant of cortical organization, it suggests that we may find functional specialization in focal cortical regions for functions that are implemented in circuits for which the neurons have to be densely connected to each other. A testable prediction of this idea is that neurons within face-selective patches of monkey cortex must be richly interconnected, either directly or via webs of inhibitory interneurons found in those same regions. A further prediction of the axon-length-minimization principle is that to the extent that readout of a neural code (by the next stage of processing) requires convergence of multiple inputs on a particular neuron, it may be easier to read out a population code represented in a focal region of cortex where those inputs can all conveniently converge on a common output neuron. In a different vein the functional significance of spatial clustering in the cortex may derive from the requirement to selectively modulate a given functional circuit by way of nonsynaptic diffusible messenger molecules that can spread a few millimeters through the cortex.

Functionally Specific Cortical Regions for Computationally Different Problems?

Of course, the classic argument for functional specialization is that efficiency can be gained by division of labor (Rueffler, Hermisson, & Wagner, 2012) when the computational requirements differ across tasks (Marr, 1982). Indeed, especially for the case of faces and places, both theoretical considerations and extensive empirical evidence suggest that different kinds of representations are extracted from these stimulus classes, and different uses are made of the resulting information (Cheng & Gallistel, 1984; Hermer & Spelke, 1996; McKone & Robbins, 2010). On the other hand, the observed clustered selectivity for faces, places, and bodies does not in itself imply qualitative differences in the computations and representations entailed in the perception of one of these categories versus another. We might have functional selectivity of the relevant neuronal responses for each category without any fundamental differences in the kinds of computations conducted for each, just as we see in retinotopic cortex, where completely nonoverlapping pools of neurons code for visual information in one visual field location versus another, but fundamentally similar computations are conducted by each. A crucial question for the enterprise of using functional specificity of the brain to infer fundamental components of the mind will therefore be: Which cortical selectivities reflect fundamentally different underlying cognitive processes, and which simply reflect convenient compartmentalization of similar processes?

CONCLUSIONS

Over the past 15 years, fMRI studies have taught us a great deal about the functional organization of the VVP in humans. We have learned that the machinery that conducts visual object recognition in humans is not a

homogeneous mass of tissue but instead a richly structured system composed of functionally distinct regions, each found in approximately the same location in every healthy subject. Exciting directions for future research will exploit a suite of powerful new methods, including the ability to relate the "representational dissimilarity matrices" extracted in each region (or the whole ventral pathway) via fMRI with representational spaces derived from behavioral and monkey single-unit data (Kriegeskorte et al., 2008) and the ability to link functionally defined regions and functional profiles with cytoarchitecture (Caspers et al., in press) and connectivity (Saygin et al., 2011).

Yet many fundamental questions have proven to be difficult or impossible to answer with current methods available in humans. What information is represented in each region at the spatial and temporal resolution of actual neural responses (Freiwald & Tsao, 2010)? What neural circuits extract this information? What is the causal role of each region in perception? What is the connectivity among these various regions, and between each of them and the rest of the brain (Moeller, Freiwald, & Tsao, 2008)? Further, how do these regions get wired up in the brain in development, and what are the relative roles of experience (Baker et al., 2007; Srihasam et al., 2012) and genes (Duchaine, Germine, & Nakayama, 2007; Sugita, 2008; Turati, Bulf, & Simion, 2008; Wilmer et al., 2010; Zhu et al., 2010) in this process? One of the most exciting developments of the last decade has been the discovery of face-selective (Tsao et al., 2006) and place-selective (Nasr et al., 2011) regions in the temporal lobes of macaques. These discoveries offer the possibility that the questions that have proven most intractable in humans will be answered by work on macaques, where it is possible to provide much richer characterizations of neural representations and their underlying circuits, the causal roles of these circuits in behavior, the structural correlates of functionally defined regions, the interplay of genes and experience that wire those regions up during development, and the interactions among those regions during task performance.

ACKNOWLEDGMENTS

The writing of this chapter was supported by EY13455, by a grant from the Ellison Medical Foundation to N.K., and by a Simons Foundation Autism Research Initiative (SFARI) postdoctoral fellowship to D.D.D.

REFERENCES

Afraz, S. R., Kiani, R., & Esteky, H. (2006). Microstimulation of inferotemporal cortex influences face categorization. *Nature, 442,* 692–695.

Aguirre, G. K., & D'Esposito, M. (1999). Topographical disorientation: A synthesis and taxonomy. *Brain, 122,* 1613–1628. doi:10.1093/brain/122.9.1613.

Aguirre, G. K., Detre, J. A., Alsop, D. C., & D'Esposito, M. (1996). The parahippocampus subserves topographical learning in man. *Cerebral Cortex, 6,* 823–829. doi:10.1093/cercor/6.6.823.

Allison, T., Puce, A., Spencer, D. D., & McCarthy, G. (1999). Electrophysiological studies of human face perception. I: Potentials generated in occipitotemporal cortex by face and non-face stimuli. *Cerebral Cortex, 9,* 415. doi:10.1093/cercor/9.5.415.

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

Baker, C. I., Hutchison, T. L., & Kanwisher, N. (2007). Does the fusiform face area contain subregions highly selective for nonfaces? *Nature Neuroscience, 10,* 3–4.

Baker, C. I., Liu, J., Wald, L. L., Kwong, K. K., Benner, T., & Kanwisher, N. (2007). Visual word processing and experiential origins of functional selectivity in human extrastriate cortex. *Proceedings of the National Academy of Sciences of the United States of America, 104,* 9087–9092. doi:10.1073/pnas.0703300104.

Bar, M. (2004). Visual objects in context. *Nature Reviews. Neuroscience, 5,* 617–629.

Bar, M., Tootell, R. B., Schacter, D. L., Greve, D. N., Fischl, B., Mendola, J. D., et al. (2001). Cortical mechanisms specific to explicit visual object recognition. *Neuron, 29,* 529–535. doi:10.1016/S0896-6273(01)00224-0.

Barlow, H. B. (1995). The neuron doctrine in perception. *Cognitive Neurosciences, 1,* 415–436.

Baumann, O., & Mattingley, J. B. (2010). Medial parietal cortex encodes perceived heading direction in humans. *Journal of Neuroscience, 30,* 12897–12901.

Bedny, M., Pascual-Leone, A., & Saxe, R. (2009). Growing up blind does not change the neural bases of theory of mind. *Proceedings of the National Academy of Sciences of the United States of America, 106,* 11312–11317. doi:10.1073/pnas.0900010106.

Bracci, S., Ietswaart, M., Peelen, M. V., & Cavina-Pratesi, C. (2010). Dissociable neural responses to hands and non-hand body parts in human left extrastriate visual cortex. *Journal of Neurophysiology, 103,* 3389–3397.

Carlin, J. D., Rowe, J. B., Kriegeskorte, N., Thompson, R., & Calder, A. J. (2011). Direction-sensitive codes for observed head turns in human superior temporal sulcus. *Cerebral Cortex,* 22, 735–744. doi:10.1093/cercor/bhr061.

Caspers, J., Zilles, K., Eickhoff, S. B., Schleicher, A., Mohlberg, H., & Amunts, K. (2013). Cytoarchitectonical analysis and probabilistic mapping of two extrastriate areas of the human posterior fusiform gyrus. *Brain Structure & Function, 218,* 511–526. doi:10.1007/s00429-012-0411-8.

Chan, A. W., Peelen, M. V., & Downing, P. E. (2004). The effect of viewpoint on body representation in the extrastriate body area. *Neuroreport, 15,* 2407–2410.

Cheng, K., & Gallistel, C. R. (1984). Testing the geometric power of an animal's spatial representation. In H. L. Roitblatt, T. G. Bever, & H. S. Terrace (Eds.), *Animal cognition* (pp. 409–423). London: Lawrence Erlbaum Associates.

Chklovskii, D. B., & Koulakov, A. A. (2004). Maps in the brain: What can we learn from them? *Annual Review of Neuroscience, 27,* 369–392.

Cohen, L., Dehaene, S., Naccache, L., Lehericy, S., Dehaene-Lambertz, G., Henaff, M. A., et al. (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*(Pt.2), 291–307. doi:10.1093/brain/123.2.291.

Conway, B. R., & Tsao, D. Y. (2009). Color-tuned neurons are spatially clustered according to color preference within alert macaque posterior inferior temporal cortex. *Proceedings of the National Academy of Sciences of the United States of America, 106,* 1–6. doi:10.1073/pnas.0810943106.

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. *Journal of Neuroscience, 31,* 11305–11312.

Dilks, D. D., Julian, J. B., Paunov, A., & Kanwisher, N. (2013). The occipital place area (OPA) is causally and selectively involved in scene perception. *Journal of Neuroscience, 33,* 1331–1336. doi:10.1523/JNEUROSCI.4081-12.2013.

Downing, P. E., Chan, A. W., Peelan, M. V., Dodds, C. M., & Kanwisher, N. (2006). Domain specificity in visual cortex. *Cerebral Cortex, 16,* 1453–1461.

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.

Downing, P. E., & Peelen, M. V. (2011). The role of occipitotemporal body-selective regions in person perception. *Cognitive Neuroscience, 2,* 186–203.

Duchaine, B., Germine, L., & Nakayama, K. (2007). Family resemblance: Ten family members with prosopagnosia and within-class object agnosia. *Cognitive Neuropsychology, 24,* 419–430.

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

Epstein, R. (2005). The cortical basis of visual scene processing. *Visual Cognition, 12,* 954–978.

Epstein, R., De Yoe, E., Press, D., & Kanwisher, N. (2001). Neuropsychological evidence for a topographical learning mechanism in parahippocampal cortex. *Cognitive Neuropsychology, 18,* 481–508.

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

Epstein, R. A., & Ward, E. J. (2010). How reliable are visual context effects in the parahippocampal place area? *Cerebral Cortex, 20,* 294–303.

Ewbank, M. P., & Andrews, T. J. (2008). Differential sensitivity for viewpoint between familiar and unfamiliar faces in human visual cortex. *NeuroImage, 40,* 1857–1870. doi:10.1016/j.neuroimage.2008.01.049.

Foldiak, P., & Young, M. P. (1995). *Handbook of brain theory and neural networks.* Cambridge, MA: MIT Press.

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

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

Furl, N., Garrido, L., Dolan, R. J., Driver, J., & Duchaine, B. (2011). Fusiform gyrus face selectivity relates to individual differences in facial recognition ability. *Journal of Cognitive Neuroscience, 23,* 1723–1740.

Ganaden, R. E., Mullin, C. R., & Steeves, J. K. E. (2013). Transcranial magnetic stimulation to the transverse occipital sulcus affects scene but not object processing. *Journal of Cognitive Neuroscience* [epub ahead of print]. doi:10.1162/jocn_a_00372.

Garrido, L., Duchaine, B., & Nakayama, K. (2008). Face detection in normal and prosopagnosic individuals. *Journal of Neuropsychology, 2*(Pt 1), 119–140.

Gauthier, I., Behrmann, M., & Tarr, M. J. (1999). Can face recognition really be dissociated from object recognition? *Journal of Cognitive Neuroscience, 11,* 349–370.

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,* 495–504.

Grill-Spector, K. (2003). The neural basis of object perception. *Current opinion in neurobiology, 13,* 159–166.

Grill-Spector, K., Kushnir, T., Edelman, S., Avidan, G., Itzchak, Y., & Malach, R. (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron, 24,* 187–203.

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

Grill-Spector, K., Kushnir, T., Hendler, T., & Malach, R. (2000). The dynamics of object-selective activation correlate with recognition performance in humans. *Nature Neuroscience, 3,* 837–843.

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

Gschwind, M., Pourtois, G., Schwartz, S., Van De Ville, D., & Vuilleumier, P. (2011). White-matter connectivity between face-responsive regions in the human brain. *Cerebral Cortex, 22,* 1564–1576. doi:10.1093/cercor/bhv226.

Habib, M., & Sirigu, A. (1987). Pure topographical disorientation: A definition and anatomical basis. *Cortex, 23,* 73–85.

Harel, A., Kravitz, D. J., & Baker, C. I. (2013). Deconstructing visual scenes in cortex: Gradients of object and spatial layout information. *Cerebral Cortex, 2,* 947–957. doi:10.1093/cercor/bhs091.

Hasson, U., Harel, M., Levy, I., & Malach, R. (2003). Large-scale mirror-symmetry organization of human occipito-temporal object areas. *Neuron, 37,* 1027–1041.

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

Haushofer, J., Livingstone, M. S., & Kanwisher, N. (2008). Multivariate patterns in object-selective cortex dissociate perceptual and physical shape similarity. *PLoS Biology, 6,* e187. doi:10.1371/journal.pbio.0060187.

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.

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

Haxby, J. V., Horwitz, B., Ungerleider, L. G., Maisog, J. M., Pietrini, P., & Grady, C. L. (1994). The functional organization of human extrastriate cortex: A PET-RCBF study of selective attention to faces and locations. *Journal of Neuroscience, 14,* 6336–6353.

Hermer, L., & Spelke, E. (1996). Modularity and development: The case of spatial reorientation. *Cognition, 61,* 195–232.

James, T. W., Culham, J. C., Humphrey, G. K., Milner, A. D., & Goodale, M. A. (2003). Ventral occipital lesions impair object recognition but not object-directed grasping: An fMRI study. *Brain, 126,* 2463–2475. doi:10.1093/brain/awg248.

Julian, J. B., Fedorenko, E., Webster, J., & Kanwisher, N. (2012). An algorithmic method for functionally defining regions of interest in the ventral visual pathway. *NeuroImage, 60,* 2357–2364. doi:10.1016/j.neuroimage.2012.02.055.

Kanwisher, N. & Barton, J. J. S. (2011). The functional architecture of the face system: Integrating evidence from fMRI and patient studies. In J. Haxby, M. Johnson, G. Rhodes, & A. Calder (Eds.), *The handbook of face perception* (pp. 111–130). Oxford: Oxford University Press.

Kanwisher, N. (2001). Neural events and perceptual awareness. *Cognition, 79,* 89–113.

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

Kanwisher, N., & Schwarzlose, R. F. (2008). *Principles governing the large-scale organization of object selectivity in ventral visual cortex.* Ph.D. Thesis, Massachusetts Institute of Technology.

Kanwisher, N., Woods, R. P., Iacoboni, M., & Mazziotta, J. C. (1997b). A locus in human extrastriate cortex for visual shape analysis. *Journal of Cognitive Neuroscience, 9,* 133–142.

Kanwisher, N., & Yovel, G. (2006). The fusiform face area: A cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 361,* 2109–2128. doi:10.1098/rstb.2006.1934.

Kim, J. G., Biederman, I., Lescroart, M. D., & Hayworth, K. J. (2009). Adaptation to objects in the lateral occipital complex (LOC): Shape or semantics? *Vision Research, 49,* 2297–2305. doi:10.1016/j.visres.2009.06.020.

Konkle, T. (2011). *The role of real-world size in object representation.* Ph.D. Thesis, Massachusetts Institute of Technology.

Kourtzi, Z., & Kanwisher, N. (2001). Representation of perceived object shape by the human lateral occipital complex. *Science, 293,* 1506–1509.

Kravitz, D. J., Peng, C. S., & Baker, C. I. (2011). Real-world scene representations in high-level visual cortex: It's the spaces more than the places. *Journal of Neuroscience, 31,* 7322–7333.

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

Lashkari, D., Sridharan, R., Vul, E., Hsieh, P. J., Kanwisher, N., & Golland, P. (2011). Search for patterns of functional specificity in the brain: A nonparametric hierarchical Bayesian model for group fMRI data. *NeuroImage, 59,* 1348–1368. doi:10.1016/j.neuroimage.2011.08.031.

Liu, J., Harris, A., & Kanwisher, N. (2010). Perception of face parts and face configurations: An fMRI study. *Journal of Cognitive Neuroscience, 22,* 203–211.

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

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information.* San Francisco: W. H. Freeman.

McCarthy, G., Puce, A., Belger, A., & Allison, T. (1999). Electrophysiological studies of human face perception. II: Response properties of face-specific potentials generated in occipitotemporal cortex. *Cerebral Cortex, 9,* 431.

McCarthy, G., Puce, A., Gore, J. C., & Allison, T. (1997). Face-specific processing in the human fusiform gyrus. *Journal of Cognitive Neuroscience, 9,* 605–610.

McKone, E. M., & Robbins, R. R. (2011). Are faces special? In J. Haxby, M. Johnson, G. Rhodes, & A Calder (Eds.), *The handbook of face perception* (pp. 149–176). Oxford: Oxford University Press.

Mendez, M. F., & Cherrier, M. M. (2003). Agnosia for scenes in topographagnosia. *Neuropsychologia, 41,* 1387–1395.

Moeller, S., Freiwald, W. A., & Tsao, D. Y. (2008). Patches with links: A unified system for processing faces in the macaque temporal lobe. *Science, 320,* 1355–1359.

Moro, V., Urgesi, C., Pernigo, S., Lanteri, P., Pazzaglia, M., & Aglioti, S. M. (2008). The neural basis of body form and body action agnosia. *Neuron, 60,* 235–246.

Mullin, C. R., & Steeves, J. K. E. (2011). TMS to lateral occipital cortex disrupts object processing but facilitates scene processing. *Journal of Cognitive Neuroscience, 23,* 4174–4184.

Nasr, S., Liu, N., Devaney, K. J., Yue, X., Rajimehr, R., Ungerleider, L. G., et al. (2011). Scene-selective cortical regions in human and nonhuman primates. *Journal of Neuroscience, 31,* 13771–13785. doi:10.1523/JNEUROSCI.2792-11.2011.

Ohki, K., Chung, S., Ch'ng, Y. H., Kara, P., & Reid, R. C. (2005). Functional imaging with cellular resolution reveals precise micro-architecture in visual cortex. *Nature, 433,* 597–603.

Ohki, K., & Reid, R. C. (2007). Specificity and randomness in the visual cortex. *Current Opinion in Neurobiology, 17,* 401–407.

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

Olshausen, B. A., & Field, D. J. (2004). Sparse coding of sensory inputs. *Current Opinion in Neurobiology, 14,* 481–487.

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. doi:10.1016/j.neuroimage.2009.11.022.

Op de Beeck, H. P., Haushofer, J., & Kanwisher, N. G. (2008). Interpreting fMRI data: Maps, modules and dimensions. *Nature Reviews. Neuroscience, 9,* 123–135.

Orlov, T., Makin, T. R., & Zohary, E. (2010). Topographic representation of the human body in the occipitotemporal cortex. *Neuron, 68,* 586–600.

O'Toole, A. J., Jiang, F., Abdi, H., & Haxby, J. V. (2005). Partially distributed representations of objects and faces in ventral temporal cortex. *Journal of Cognitive Neuroscience, 17,* 580–590.

Park, S., Brady, T. F., Greene, M. R., & Oliva, A. (2011). Disentangling scene content from spatial boundary: Complementary roles for the parahippocampal place area and lateral occipital complex in representing real-world scenes. *Journal of Neuroscience, 31,* 1333–1340.

Park, S., & Chun, M. M. (2009). Different roles of the parahippocampal place area (PPA) and retrosplenial cortex (RSC) in panoramic scene perception. *NeuroImage, 47,* 1747–1756. doi:10.1016/j.neuroimage.2009.04.058.

Peelen, M. V., & Downing, P. E. (2005). Within-subject reproducibility of category-specific visual activation with functional MRI. *Human Brain Mapping, 25,* 402–408.

Phillips, M. L., & David, A. S. (1997). Viewing strategies for simple and chimeric faces: An investigation of perceptual bias in normals and schizophrenic patients using scan paths. *Brain and Cognition, 35,* 225–238.

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,* 319–324.

Pitcher, D., Dilks, D. D., Saxe, R. R., Triantafyllou, C., & Kanwisher, N. (2011). Differential selectivity for dynamic versus static information in face-selective cortical regions. *NeuroImage, 56,* 2356–2363. doi:10.1016/j.neuroimage.2011.03.067.

Pitcher, D., Walsh, V., Yovel, G., & Duchaine, B. (2007). TMS evidence for the involvement of the right occipital face area in early face processing. *Current Biology, 17,* 1568–1573. doi:10.1016/j.cub.2007.07.063.

Potter, M. C. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology. Human Learning and Memory, 5,* 509–522.

Puce, A., Allison, T., Asgari, M., Gore, J. C., & McCarthy, G. (1996). Differential sensitivity of human visual cortex to faces, letterstrings, and textures: A functional magnetic resonance imaging study. *Journal of Neuroscience, 16,* 5205–5215.

Puce, A., Allison, T., Bentin, S., Gore, J. C., & McCarthy, G. (1998). Temporal cortex activation in humans viewing eye and mouth movements. *Journal of Neuroscience, 18,* 2188–2199.

Puce, A., Allison, T., Gore, J. C., & McCarthy, G. (1995). Face-sensitive regions in human extrastriate cortex studied by functional MRI. *Journal of Neurophysiology, 74,* 1192–1199.

Puce, A., Allison, T., & McCarthy, G. (1999). Electrophysiological studies of human face perception. III: Effects of top-down processing on face-specific potentials. *Cerebral Cortex, 9,* 445.

Rajimehr, R., Devaney, K. J., Bilenko, N. Y., Young, J. C., & Tootell, R. B. (2011). The "parahippocampal place area" responds preferentially to high spatial frequencies in humans and monkeys. *PLoS Biology, 9,* e1000608.

Rajimehr, R., Young, J. C., & Tootell, R. B. (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,* 1995–2000. doi:10.1073/pnas.0807304106.

Reddy, L., & Kanwisher, N. (2007). Category selectivity in the ventral visual pathway confers robustness to clutter and diverted attention. *Current Biology, 17,* 2067–2072. doi:10.1016/j.cub.2007.10.043.

Rossion, B., Hanseeuw, B., & Dricot, L. (2012). Defining face perception areas in the human brain: A large-scale factorial fMRI face localizer analysis. *Brain and Cognition, 79,* 138–157.

Rueffler, C., Hermisson, J., & Wagner, G. P. (2012). Evolution of functional specialization and division of labor. *Proceedings of the National Academy of Sciences of the United States of America, 109,* E326–E335. doi:10.1073/pnas.1110521109.

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

Saygin, Z. M., Osher, D. E., Koldewyn, K., Reynolds, G., Gabrieli, J. D. E., & Saxe, R. R. (2011). Anatomical connectivity patterns predict face selectivity in the fusiform gyrus. *Nature Neuroscience, 15,* 321–327.

Schiltz, C., & Rossion, B. (2006). Faces are represented holistically in the human occipito-temporal cortex. *NeuroImage, 32,* 1385–1394.

Schwarzlose, R. F., Baker, C. I., & Kanwisher, N. (2005). Separate face and body selectivity on the fusiform gyrus. *Journal of Neuroscience, 25,* 11055–11059.

Schwarzlose, R. F., Swisher, J. D., Dang, S., & Kanwisher, N. (2008). The distribution of category and location information across object-selective regions in human visual cortex. *Proceedings of the National Academy of Sciences of the United States of America, 105,* 4447–4452. doi:10.1073/pnas.0800431105.

Sergent, J., & Signoret, J. L. (1992). Implicit access to knowledge derived from unrecognized faces in prosopagnosia. *Cerebral Cortex, 2,* 389–400.

Silvanto, J., Schwarzkopf, D. S., Gilaie-Dotan, S., & Rees, G. (2010). Differing causal roles for lateral occipital cortex and occipital face area in invariant shape recognition. *European Journal of Neuroscience, 32,* 165–171.

Srihasam, K., Mandeville, J. B., Morocz, I. A., Sullivan, K. J., & Livingstone, M. S. (2012). Behavioral and anatomical consequences of early versus late symbol training in macaques. *Neuron, 73,* 608–619.

Stettler, D. D., & Axel, R. (2009). Representations of odor in the piriform cortex. *Neuron, 63,* 854–864.

Sugita, Y. (2008). Face perception in monkeys reared with no exposure to faces. *Proceedings of the National Academy of Sciences of the United States of America, 105,* 394–398.

Takahashi, N., Kawamura, M., Shiota, J., Kasahata, N., & Hirayama, K. (1997). Pure topographic disorientation due to right retrosplenial lesion. *Neurology, 49,* 464–469.

Tarr, M. J., & Gauthier, I. (2000). FFA: A flexible fusiform area for subordinate-level visual processing automatized by expertise. *Nature Neuroscience, 3,* 764–769.

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,* 1467–1477.

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature, 381,* 520–522.

Tian, L., Jiang, T., Liu, Y., Yu, C., Wang, K., Zhou, Y., et al. (2007). The relationship within and between the extrinsic and intrinsic systems indicated by resting state correlational patterns of sensory cortices. *NeuroImage, 36,* 684–690. doi:10.1016/j.neuroimage.2007.03.044.

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

Turati, C., Bulf, H., & Simion, F. (2008). Newborns' face recognition over changes in viewpoint. *Cognition, 106,* 1300–1321. doi:10.1016/j.cognition.2007.06.005.

Turk-Browne, N. B., Norman-Haignere, S. V., & McCarthy, G. (2010). Face-specific resting functional connectivity between the fusiform gyrus and posterior superior temporal sulcus. *Frontiers in Human Neuroscience, 4,* 176. doi:10.3389/fnhum.2010.00176.

Urgesi, C., Berlucchi, G., & Aglioti, S. M. (2004). Magnetic stimulation of extrastriate body area impairs visual processing of nonfacial body parts. *Current Biology, 14,* 2130–2134. doi:10.1016/j.cub.2004.11.031.

van Koningsbruggen, M. G., Peelen, M. V., & Downing, P. E. (2013). A causal role for the extrastriate body area in detecting people in real-world scenes. *Journal of Neuroscience, 33,* 7003–7010. doi:10.1523/JNEUROSCI.2853-12.2013.

Wada, Y., & Yamamoto, T. (2001). Selective impairment of facial recognition due to a haematoma restricted to the right fusiform and lateral occipital region. *Journal of Neurology, Neurosurgery, and Psychiatry, 71,* 254–257.

Weiner, K. S., & Grill-Spector, K. (2011). Neural representations of faces and limbs neighbor in human high-level visual cortex: Evidence for a new organization principle. *Psychological Research* [epub ahead of print December 3, 2011].

Weiner, K. S., & Grill-Spector, K. (2012). The improbable simplicity of the fusiform face area. *Trends in Cognitive Sciences, 16,* 251–254. doi:10.1016/j.tics.2012.03.003.

Williams, M. A., Berberovic, N., & Mattingley, J. B. (2007). Abnormal FMRI adaptation to unfamiliar faces in a case of developmental prosopamnesia. *Current Biology, 17,* 1259–1264. doi:10.1016/j.cub.2007.06.042.

Wilmer, J. B., Germine, L., Chabris, C. F., Chatterjee, G., Williams, M., Loken, E., et al. (2010). Human face recognition ability is specific and highly heritable. *Proceedings of the National Academy of Sciences of the United States of America, 107,* 5238–5241. doi:10.1073/pnas.0913053107.

Winston, J. S., Henson, R. N., 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,* 1830–1839.

Yovel, G., & Kanwisher, N. (2004). Face perception domain specific, not process specific. *Neuron, 44,* 889–898.

Yovel, G., & Kanwisher, N. (2005). The neural basis of the behavioral face-inversion effect. *Current Biology, 15,* 2256–2262. doi:10.1016/j.cub.2005.10.072.

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. *Cerebral Cortex, 21,* 35–47. doi:10.1093/cercor/bhq050.

Zhu, Q., Song, Y., Hu, S., Li, X., Tian, M., Zhen, Z., et al. (2010). Heritability of the specific cognitive ability of face perception. *Current Biology, 20,* 137–142. doi:10.1016/j.cub.2009.11.067.