21

WHAT MAKES US HUMAN? FACE IDENTITY RECOGNITION

Bruno Rossion

Introduction

The human brain is an organ weighing 1.2–1.5 kg that, together with the spinal cord, forms the central nervous system. It contributes to regulating body functions, processing sensory inputs of the environment, and controlling bodily actions. It is widely considered to be the most complex biological organ of the human body, consisting of about 86 billion nerve cells, or neurons, which present an astonishing level of morphological and functional variability. Each neuron is connected to other neurons by tens of thousands of synaptic connections, making this system extremely complex and difficult to understand. While the human brain is the product of a long biological evolution, it differs considerably from the brain of other animal species at multiple micro- and macroscopic levels of anatomo-functional organization (Bryant & Preuss, 2018; Gazzaniga, 2009; Herculano-Houzel, 2016; Kaas, 2019).

In its interaction with the environment, the brain can essentially be considered as a biological organ of recognition (Edelman, 2004): its primary function is to *discriminate* a myriad of sensory stimuli coming from a diverse, inherently ambiguous, and dynamic environment, and *to generalize* (or reproduce) these discriminative/selective responses across various instances of these stimuli based on their physical properties, as well as the brain's own biological constraints and past experience. For example, the adult human brain is able to recognize a wide range of stimuli as edible fruits, or different instances of insects as dangerous or harmless, and behave adaptively when facing them in daily life.

In psychology, the term "recognition" usually implies a judgment of previous occurrence (specifically "the ability to identify information as having been encountered before," according to the *APA Dictionary of Psychology*; see also Mandler, 1980). This is problematic if only because it confines recognition to experience, ruling out the contribution of phylogenetic memory. With such a restrictive semantic definition, *recognition* usually refers to a function different from *categorization* – two terms that are used preferentially by relatively independent scientific communities in psychology and cognitive (neuro)science (Palmeri & Gauthier, 2004). Here, I adopt a more general biological perspective, considering that recognition is essentially a categorization function that, in human adults, is largely automatic and associated with different levels of consciousness (Rossion & Retter, 2020). Without this recognition/categorization function (i.e., if each stimulus were treated as unique at every moment in time), we would be overwhelmed by the amount of sensory inputs received,

DOI: 10.4324/9781003051817-26

and lost in our interaction with the world. Hence, recognition/categorization is arguably the most important brain function, providing a key platform for learning, memory, attention, decision-making, language, and social communication, among other faculties (Murphy, 2002).

Inputs from the environment are dealt with by various sensory receptors in biological organisms, with vision being largely considered in the scientific community as the most elaborate sensory modality in humans (and primates in general). Visual stimuli vary substantially across time and space, with variations of luminance and wavelength providing extremely rich inputs. In humans and other primates, considerable brain portions deal mainly, or only, with visual inputs (Felleman & Van Essen, 1991) and the experimental studies in cognitive (neuro)science devoted to vision largely outnumber studies performed with stimuli suited to other sensory modalities (Hutmacher, 2019).

Among all sensory inputs to the human visual system, human faces constitute arguably the most important in our environment: they convey many cues for social communication and are frequently encountered, especially in modern societies and in the age of (social) media. Further, a face is a complex stimulus, containing multiple nameable parts (e.g., eyes, mouth, nose) and internal sources of physical variations interpreted in terms of shape, texture, and color. The face is thus extremely well suited for investigating the most refined mechanisms of recognition. In particular, often considered as the quintessential whole, or *Gestalt*, the face appears as the ideal stimulus to understand the complex relationships between wholes and parts in (visual) recognition (Rossion, 2013). For all these reasons, the human face undoubtedly holds a special place in scientific research, having become possibly the most widely studied stimulus in cognitive (neuro)science (Calder et al., 2011).

The key point of this chapter is that face recognition is specific to the human species in terms of expertise and neural mechanisms. This is not only because a human face is a powerful nonverbal stimulus communicating meaning, but more fundamentally because its recognition based on visual inputs only, in the adult human brain, originates from and is built upon multimodal (including verbal) associations. The recognition of identity from the face, face identity recognition (FIR), is used here as the primary example to illustrate this view. Since facial identity is not an objective conveyor of information to decode, but rather a sign to interpret based on semantic knowledge, understanding the biological bases of FIR is important for the agenda of neurosemiotics, and a neurosemiotic perspective can inspire future research on this topic.

Following this brief introduction, the chapter is divided into five further sections to first introduce the face as a stimulus, followed by insights on FIR, its development, and its fundamental difference between humans and other animal species. The next section will describe the human-specific neural circuitry subtending this function, highlighting key differences in cortical face networks with the best available animal model in cognitive neuroscience, the macaque brain. Following this analysis, an original model for the developmental origin and structuring of human cortical memories of face identities based on reentrant semantically driven identity representations will briefly be proposed before concluding remarks are made on this neurosemiotic perspective of human face (identity) recognition.

The human face as a powerful sign of social communication

Faces emerged through evolution in animal species over half a billion years ago as a biological interface between the central nervous system and the physical environment (McNeill, 2000). From a biophysical perspective, a face is a collection of sensory receptors – typically the eyes, nose, mouth, and ears – positioned together at the front of the body and in proximity to the brain. Importantly for an animal's survival, this arrangement enables information from the outside world to be reacted to appropriately and rapidly (Rossion & Retter, 2020). Many animal species have a face, with the sensory receptors organized in an overall quasi-symmetric configuration that often bears a striking resemblance to the human face (Bruce & Young, 1998). Despite this similarity, the human face has many characteristics that make it unique, even when compared with our closest cousin, the chimpanzee: a large head size relative to the body, a flat face due to the development of the prefrontal cortex, tooth size reduction, a prominent nose, a chin, less facial hair, thinner eyebrows, and a relatively small distance between the nose and the eyes.

Since populations of modern humans – *Homo Sapiens* – started to leave Africa around 100,000 years ago and settle on different continents, their body morphology, including the face, transformed in response to different natural selection pressures – that is, different physical and social environments (Ruff, 2002). Consequently, despite the extremely high genetic similarity between different modern human populations (Cavalli-Sforza & Cavalli-Sforza, 1995), there are important and markedly visible differences between their faces, both in terms of the bone structure and the reflectance of light on the skin (Bruce & Young, 1998). For instance, sub-Saharan Africans have darker skin and wider, more bulbous noses than Asians or Western Europeans, with the latter being characterized by thinner, pointier noses (Farkas, 1994; Zaidi et al., 2017).

In modern human populations, the face evolved to have primary importance for social functioning, conveying both quickly changeable cues to recognize/categorize emotion from the expression, the direction of attention from the position of the head and gaze, and the state of health from the skin tone – as well as relatively invariant cues to recognize/categorize age, sex, race, attractiveness, and personality traits such as dominance and trustworthiness (Bruce & Young, 1998; Todorov, 2017). The face is thus a multifaceted, ecologically valid construct with extreme importance for human social life. In this sense, the human face is an extremely powerful sign conveying meaning to the observer, and should be a primary object of study of (neuro)semiotics.

Face identity recognition: A human-specific expertise

The challenge and quality of human face identity recognition

Recognizing a face as belonging to a specific identity, or FIR, is particularly challenging in humans. This difficulty follows from two main factors. First, while individual faces certainly differ more in humans than in other animal species (Sheehan & Nachman, 2014), all human faces, particularly within a genetically homogenous group, share similar features and their overall configuration, so FIR requires fine-grained visual discrimination processes. Second, the same face identity can vary substantially under different viewing conditions (Figure 21.1), to the point where two views of the same identity may often differ physically more than two different facial identities (Burton et al., 2016).

A third factor, usually neglected, is that in most modern human societies, the number of different people encountered (in real life or through the media) is extremely large and variable over time. For this reason, the number of *categories* of human facial identities is, if not infinite, extremely large and undetermined. Hence, contrary to other face recognition functions (e.g., emotional expression, sex, age), for which there is a small and relatively stable number of categorizations to make, FIR is a wide-open flexible recognition function, making it extremely challenging. This challenge explains why there is so much natural interindividual variability in this ability in the normal population (Wilmer, 2017), and why this ability is easily lost in many neurological, neurodevelopmental and psychiatric disorders (Barton et al., 2021).

These three factors also explain why, despite major investments in technology and substantial progress over the past decades, artificial algorithms for FIR still struggle to come close to human performance and can be fooled relatively easily.[1] Since early 2020 and the COVID-19 pandemic, artificial FIR has also struggled to surmount the new challenges posed by face masks.[2]

Despite this challenge, in humans identity recognition is based primarily on the face, which among body parts carries by far the largest morphological and genetic diversity (Sheehan & Nachman, 2014). Young adults in modern societies are able to accurately identify thousands of faces on average (Jenkins et al., 2018), rapidly – that is, at a glance, and within a few hundreds of milliseconds of processing (Hsiao et al., 2008; Yan & Rossion, 2020) – and largely automatically (i.e., without volition). Thanks to these characteristics (accuracy, speed, automaticity, large and variable number of exemplars

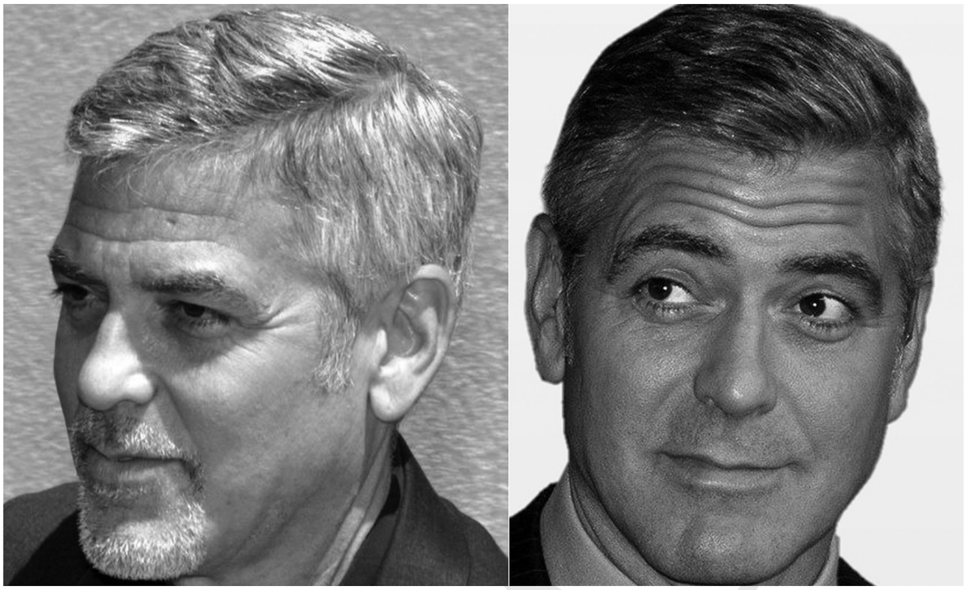

Figure 21.1 A face identity recognition (FIR) as categorization

FIR requires producing a **selective** response to a view of a face (e.g., George Clooney) compared with other (familiar or unfamiliar) face identities and to *generalize* this response across widely different signals – for example, very different views of George Clooney's face as in the display. Due to large changes in viewing conditions but also facial appearance during the lifetime, the objective physical difference can be much smaller between one view of a face identity and another facial identity (e.g., Clooney and a Clooney lookalike) than between two views of the same face identity. According to the neurosemiotic view advocated here, generalization across different views of a face identity is not based on invariant unimodal visual representations derived from an objective analysis of facial features, but is necessarily rooted in multimodal semantic memory. Pictures licensed under the Creative Commons Attribution – Share Alike 3.0 Unported. Attribution: Georges Biard. See e-book for a full-color version of this figure.

to recognize), neurotypical human adults can be considered genuine *experts* at FIR (Carey, 1992; Rossion, 2018a).

Expertise in FIR is thought to be based on the ability of the neurotypical adult brain to recognize a face identity holistically/configurally – that is, as a single unit rather than a collection of independent features (Tanaka & Farah, 1993). This recognition mode is well illustrated by visual phenomena such as the composite face illusion, in which identical top face halves aligned with differing bottom face halves, produce the vivid impression of different top face halves (Figure 21.2). This visual illusion provides compelling evidence that the recognition of the identity of a face part is mandatorily affected by the identity of the other parts (Rossion, 2013; Tanaka & Farah, 1993; Young et al., 1987) and that FIR goes well beyond the stimulus given. The importance of holistic face identity recognition is supported, among other observations (e.g., Van Belle et al., 2010), by the disruption of the composite face illusion following stimulus inversion (Rossion, 2013; Young et al., 1987; Figure 21.2) and in cases of prosopagnosia (Ramon et al., 2010), a sudden, massive, and category-selective impairment of FIR following brain damage (Bodamer, 1947; for a recent review, see Rossion, 2018b).

Development, genetics, and experience

To what extent is human FIR genetically predetermined and to what extent does it arise due to visual experience? On the one hand, developmental difficulties at FIR – most often called "developmental or congenital prosopagnosia" (Behrmann & Avidan, 2005), but more correctly "prosopdysgnosia"

Figure 21.2 The composite face illusion

This powerful illusion shows that FIR goes beyond the stimulus given. A. All five top halves (above the thin line) are physically identical. Yet, when they are aligned with distinct bottom halves (all of different identities, neutral expression, taken under the same lighting conditions), they are perceived as being different. B. If the bottom halves differ but are spatially misaligned with the top halves, one has no difficulties in perceiving the top face halves as being indeed identical. C. The same as in A but the display has been vertically flipped, showing inverted faces. All 5 "top" halves (here at the bottom of the display below the thin line) are physically identical and one does not see them as being different. See e-book for a full-color version of this figure.

(see Rossion, 2018c) – may run in families (Susilo & Duchaine, 2013), and monozygotic twins' performance at FIR tasks are more strongly correlated than dizygotic twins' performance (Wilmer et al., 2010). These observations have led to claims that genetics largely determine FIR ability. On the other hand, early visual experience appears to be critical for the development of FIR based on holistic representations (Le Grand et al., 2004) and also leads to the fine-tuning of FIR ability; at nine months of age, infants show lower FIR performance for non-experienced than frequently experienced faces – for example, other-species faces (Pascalis et al., 2002) and other-race human faces (Kelly et al., 2007). Nevertheless, even after nine years of age a predominant exposure to other-race faces can reverse such an early bias in FIR (Sangrigoli & de Schonen, 2004). Thus, experience appears to be necessary and influential for FIR*,* which continues to improve during development until adulthood (Carey, 1992). While some authors consider that this long development is essentially explained by the maturation of general cognitive factors (Crookes & McKone, 2009), there is substantial behavioral and neural evidence for a progressive, specific improvement of FIR. For instance, the face inversion effect, measured as the difference in performance for recognizing the identity of upright versus inverted faces, emerges only at a few years of age and increases until adulthood (Hills & Lewis, 2018). Moreover, the spatial extent and degree of face-selectivity in the human ventral occipito-temporal cortex, as described in the next section, continues to develop past age seven into adulthood (Golarai et al., 2017).

Importantly, although the notion of expertise has been around for decades in the field of human face recognition (Diamond & Carey, 1986; Carey, 1992), it is generally conflated with the view that expertise necessarily arises from visual experience only, and that it could be developed in a similar way for virtually any kind of non-face stimuli learned during development or even in adulthood (Diamond & Carey, 1986; Gauthier & Tarr, 1997). In this vein, "acquired" domain-general expertise has often been radically opposed to "innate" domain-specific (i.e., only for faces) expertise (McKone et al., 2007). Instead, current evidence suggests that, despite a large amount of interindividual variability in FIR ability (Wilmer et al., 2010), neurotypical human adults possess a domain-specific expertise in FIR, which emerges both from genetic factors and experience. The level of expertise that humans reach in FIR appears particularly salient when compared with FIR performance of other animal species.

A human-specific expertise in FIR

Animal species other than humans undoubtedly rely on faces of conspecifics for social communication, with the face of non-human primates, in particular, conveying important cues for recognizing emotional expression and attention from gaze and head direction (e.g., Emery et al., 1997). Yet whether the face is used as a (primary) cue for identity recognition in non-human species is much more debatable. Behavioral studies have shown that many non-human species, such as sheep (Peirce et al., 2000), fish (Wang & Takeuchi, 2017), and even wasps (Sheehan & Tibbetts, 2011) can discriminate pictures of faces of different conspecifics, drawing parallels to human performance. However, FIR performance of non-human species often follows long operant-conditioning procedures and, despite the use of a small number of individual exemplars (with often the exact same images repeated extensively), remains relatively modest (Rossion & Taubert, 2019; Figure 21.3A). Moreover, successful performance on such tasks appears to be based on low-level image characteristics (with limited across-image generalization) and/or external traits of the face – for example, antennae for wasps or ears for sheep; see respective references above (Figure 21.3A).

Importantly, this limited ability also holds for the species often considered as offering the best available animal model for the human visual system (and even for complex visual object recognition – see DiCarlo et al., 2012): the macaque monkey. Interestingly, macaques are poor at behavioral FIR tasks, requiring hundreds of trials to reach acceptable levels of performance on simple image discrimination tasks, with little generalization over new images (Parr et al., 2008) (Figure 21.3B). This is also the case for other monkey species, such as Capuchin monkeys (Pokorny & de Waal, 2009; see

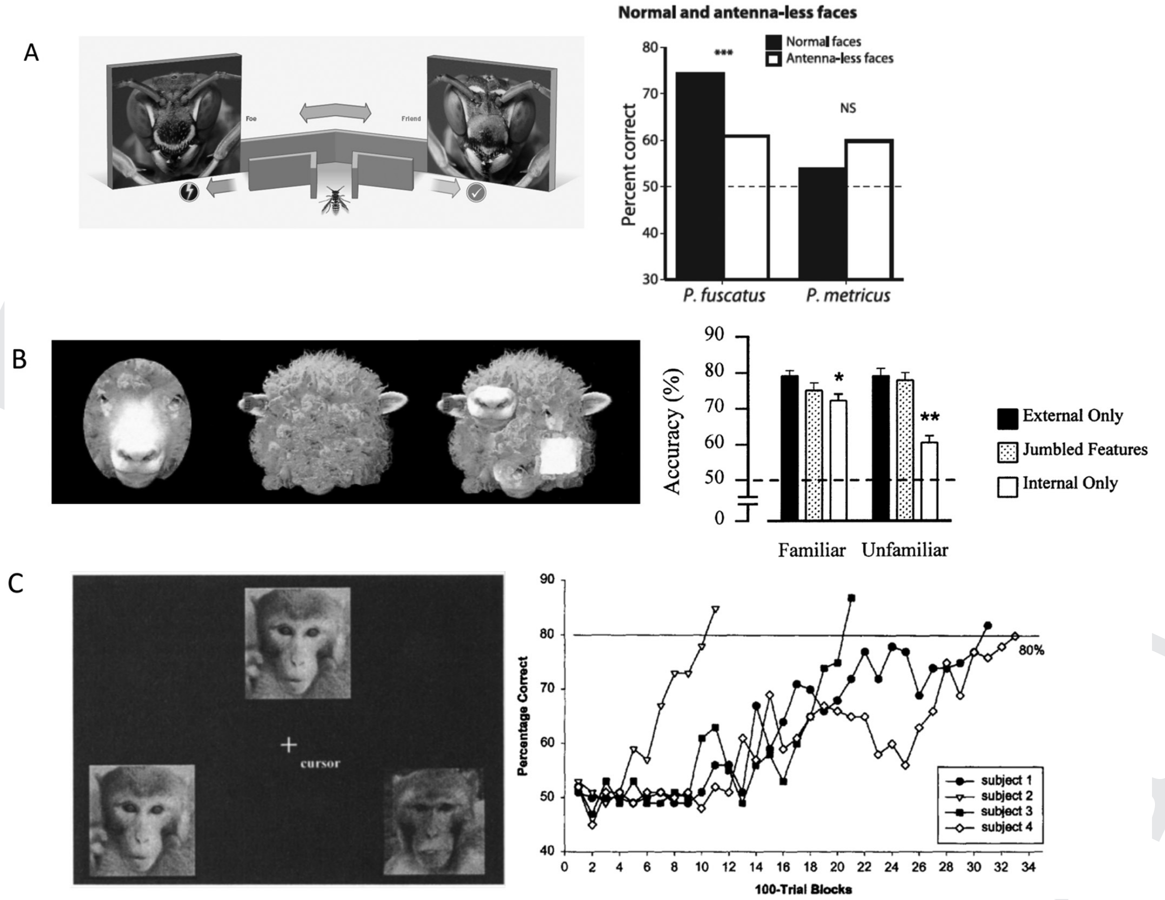

Figure 21.3 Many animal species can be trained to discriminate the identity of conspecifics from their faces

A. Polistes fuscatus wasps as shown in in A, whose recognition performance depends heavily on the presence of the antenna (from Sheehan & Tibbetts, 2011; with permission). B. Similarly, sheep can be trained to discriminate images of conspecifics, but with a

Figure 21.3 (Continued)

limited performance that, unlike in humans, is based on piecemeal information, especially external features (from Peirce et al., 2000, with permission). C. Left. Example of a sequential 2 alternative forced-choice or match-to-sample task used in the studies of Parr and colleagues (Parr et al., 2000, 2008) in which a target face is presented on top, and disappears when the monkey contacts it with a cursor driven by a joystick. After the target face is contacted, it is cleared from the monitor, and the two comparison stimuli appear below. In some versions of the task, the target face remains on the screen. Note that the two images to match are strictly identical. Right. Results of the four macaques tested at this trivial task with blocks of 100 trials, showing that they need between 1100 and 3300 trials to reach criterion (>80% performance). Introducing a new image set after reaching criterion leads to a significant drop of performance, even though the task still requires a simple image comparison (from Parr et al., 2008, with permission under a PMC Creative Commons License). See e-book for a full-color version of this figure.

the extensive review of Rossion & Taubert, 2019). Importantly, since monkeys' performance drops after the introduction of new images, this modest performance is not due to difficulties at specific task acquisition. Moreover, monkeys do not rely on qualitatively similar processes to humans for FIR. That is, they rely primarily on external features (Parr et al., 2008), and generally fail to show a decrease of performance with stimulus (picture-plane) inversion (Bruce, 1982; Griffin, 2020; Rossion & Taubert, 2019), one of the most compelling effects in human studies (Rossion, 2008). Moreover, unlike humans (e.g., Bruce et al., 2001), they do not perform better for matching pictures of familiar than unfamiliar faces for their identity (Micheletta et al., 2015; Pokorny and de Waal, 2009).

Overall, these observations cast strong doubt as to the appropriateness of the (macaque) monkey as a model for understanding the neural basis of human FIR. Although behavioral studies suggest that chimpanzees, and perhaps other apes, are better than macaques at FIR and might rely on more qualitatively similar processes to humans (Taubert et al., 2012), humans may be the only animal species living in a large social group to naturally use the face as the primary source of information for FIR (as opposed to smell, vocalization, non-facial body cues and so on), reaching expertise levels that appear unparalleled in the animal world.

Face identity recognition: A human-specific neural circuitry

What are the neural mechanisms supporting human FIR? For the reasons presented above, answering this question would go a long way towards understanding recognition/categorization, arguably the most fundamental brain function. As we shall see below, the answer to this question is also important for the agenda of neurosemiotics, which should be concerned not only with how semantic knowledge about people's identity is associated with visual constructs in the human brain, but by how semantic knowledge constrains and even generates face recognition based on visual inputs.

A right-lateralized network of face-selective regions

The neural bases of human FIR were first hinted at in patients presenting prosopagnosia – that is, after lesions to the ventral occipito-temporal cortex (VOTC), in particular the right fusiform gyrus (Cohen et al., 2019; Meadows, 1974). Since the mid-1990s, functional magnetic resonance imaging (fMRI) studies have identified clusters of voxels (of a few cubic millimeters) responding significantly more to pictures of faces than nonface objects, especially along the VOTC (in particular, the lateral fusiform gyrus) and the superior temporal sulcus (STS) (Figure 21.4) – for reviews and models, see Duchaine & Yovel (2015), Grill-Spector et al. (2017), and Haxby et al. (2000). While the human VOTC network has been associated with the recognition of relatively stable aspects of a face, in particular its identity, the more dorsal STS network is thought to be involved in the recognition of rapidly changing (i.e., dynamic) aspects of faces, such as eye gaze direction, head orientation, and facial expression (Allison et al., 2000; Bernstein & Yovel, 2015).

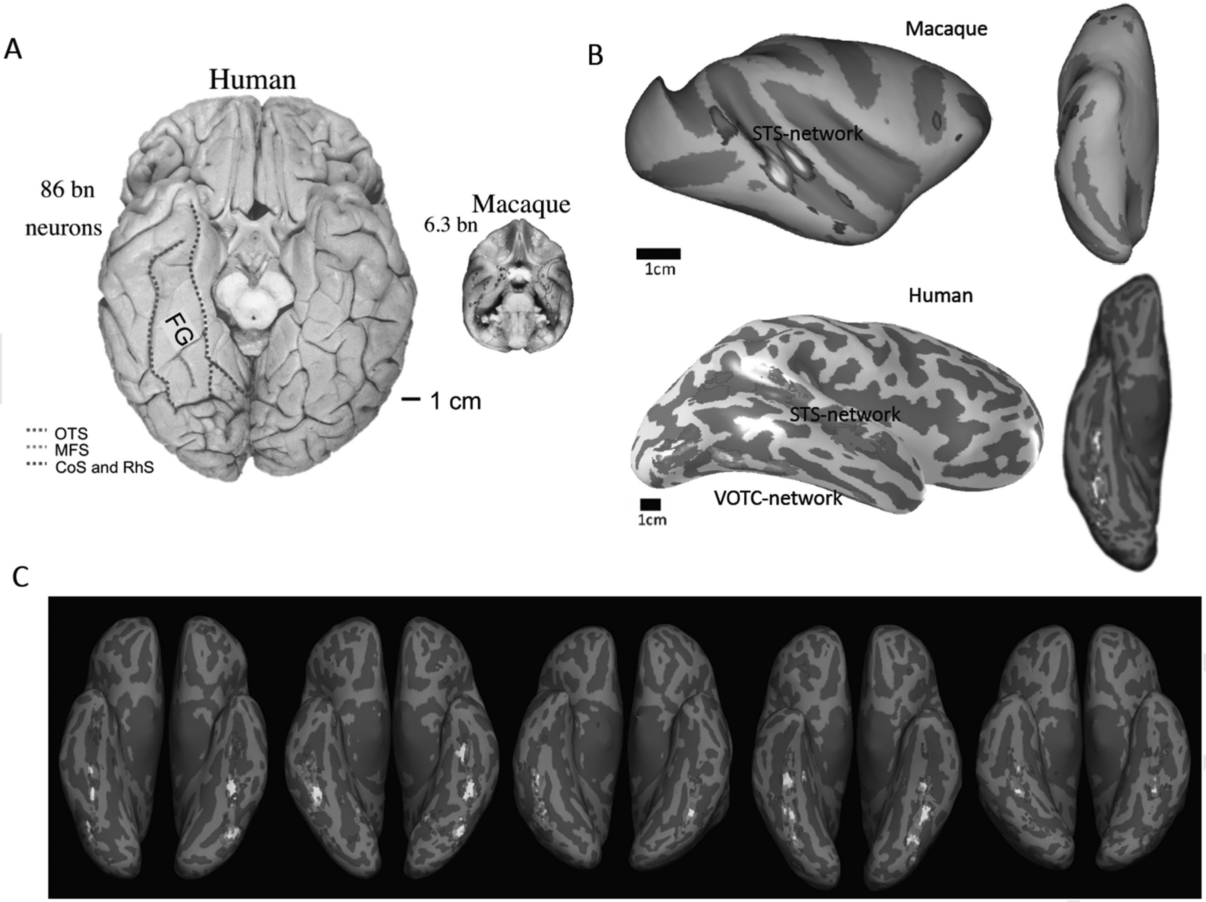

Figure 21.4 Human specificity of neural circuits for face identity recognition

A. Ventral view of a human and a macaque brain at relative sizes. The human brain contains approximately 86,000 million neurons, about 13.5 times more than a macaque brain (6376 million) (Herculano-Houzel, 2016). In humans, the key neural structures for face identity recognition run from

Figure 21.4 (Continued)

the lateral section of the inferior occipital gyrus (IOG) to the posterior (pFG), middle lateral (lFG) and anterior fusiform gyrus (aFG) up to the temporal pole. In comparison, the posterior ventral occipito-temporal cortex of the macaque brain is limited to one sulcus, with little gyrification, and no fusiform gyrus (Bryant & Preuss, 2018). B. Inflated segmented brains showing typical locations of posterior face-selective regions in macaques and humans (adapted from Weiner & Grill-Spector, 2015; with permission under a PMC Creative Commons License). In macaques, these regions are essentially found in the STS, whereas humans have regions in both the STS and the VOTC (1, 2, 3, 4 on the figure). C. Ventral surface of individual human brains, showing the wide variability across five individuals in terms of the number and size of face-selective responses (data unsmoothed, obtained here with a highly sensitive and specific fMRI face localizer, Gao et al., 2018). This ventral network of face-selective regions is found primarily in the lateral fusiform gyrus and does not exist in monkeys. See e-book for a full-color version of this figure.

Neurofunctional models of human face recognition have been generated, with modular roles attributed to clusters of voxels (usually referred to as face-selective areas) in the VOTC and STS (Duchaine & Yovel, 2015; Haxby et al., 2000) (see Figure 21.4). The face-selective cluster in the lateral middle fusiform gyrus, known as fusiform face area (FFA; Kanwisher et al., 1997) has been divided into several subclusters, such as the pFus-faces/FFA-1 and mFus-faces/FFA-2, which have been linked to cytoarchitectonically distinct regions (Grill-Spector et al., 2017).

In recent years, there has been a rise in fMRI studies using multivariate pattern analysis to "decode" the representations of (unfamiliar) individual faces in the cortical face network in order to localize and better define the neural bases of FIR. Unfortunately, these studies often report weak and inconsistent effects essentially driven by image-based cues (Nestor et al., 2011; see Rossion, 2014). In comparison, fMR-adaptation studies (Grill-Spector & Malach, 2001), which are based on reduced fMRI signal to repeated facial identities as compared with different stimuli, have reported more consistent effects in face-selective VOTC regions (Davies-Thompson et al., 2009; Gauthier et al., 2000).

However, such effects have not been reported in more anterior regions of the VOTC, which are affected by large magnetic susceptibility artifacts in fMRI (Rossion et al., 2018; Wandell, 2011). This drop in signal-to-noise ratio creates a "gap" between posterior VOTC regions and more anterior regions in the temporal pole (e.g., between regions 3 and 4 in Figure 21.4B), this latter region being linked to multimodal semantic associations, in particular knowledge about people (Collins et al., 2016; Gainotti, 2013; Lambon-Ralph et al., 2010). Interestingly, at the functional level, this gap caused by methodological factors corresponds to the advocated dissociation between "perception" and "semantic memory" for faces (Duchaine & Yovel, 2015; Haxby et al., 2000). However, lesion analyses of prosopagnosia (Cohen et al., 2019) as well as focal electrical intracerebral recordings (Jacques et al., 2020; Jonas et al., 2016) and direct stimulation (Jonas & Rossion, 2021), point to a key role of populations of neurons throughout the anterior ventral temporal cortex, rather suggesting a continuous cortical representation of faces along the whole VOTC.

The lack of a ventral face-selective network in macaque monkeys

Since there is no reason to expect, and no evidence whatsoever, that each facial identity would be uniquely associated with a specific brain region or even a neuronal population at the level of a single voxel in fMRI (containing millions of neurons; Logothetis, 2008), it is generally assumed that a full understanding of the neural mechanisms of human FIR requires complementary neural measures at finer spatial scales of brain organization, all the way down to the level of single neurons. This is where rhesus monkeys have served as a non-human primate model, potentially bridging what we learn from human lesion studies and neuroimaging to activity at the level of the single cell.

Single neurons responding selectively to faces were found in the macaque temporal cortex originally by Gross et al. (1972). Further single cell studies performed during the 1980s reported large numbers of these face-selective neurons, describing their response properties in detail (e.g., Perrett et al., 1982; see Desimone, 1991 for review). Since early studies reported variability in the firing rate of single neurons in response to images of different individual faces, this variability has been proposed as a mechanism for FIR (Rolls, 1992). Specifically, each face-selective neuron of a small population does not respond only to one face, but has a different relative response, in terms of the number of spikes/second, to each of the members of a set of facial identities. Thus, the firing rate across a population of such neurons can convey information about different individuals – that is, a distributed neural code (Rolls, 1992). Importantly, this information is thought to relate directly to physical properties of the stimuli: faces that are physically similar/different appear to be represented by similar/different distributed patterns of spike rates across the neuronal population – a "face space" (Leopold et al., 2006; Young & Yamane, 1992). In recent years, this seminal work has been extended to target, and compare, the coding of different facial identities in fMRI-defined face-selective regions of the macaque brain (Tsao et al., 2006). For instance, an influential study reported that the firing rates of approximately 200 neurons sampled from face-selective regions in two macaque brains could be used to decode visual similarity among 2000 humanoid pictures of faces and claimed to have finally cracked the neural code that allows the human brain to perform FIR (Chang & Tsao, 2017; see also Freiwald, 2020; Hesse & Tsao, 2020).

However, this claim is unsubstantiated (Rossion & Taubert, 2017), and the view that variable spike rates in a small population of neurons in the monkey brain offer a mechanistic account of human FIR is fatally flawed for several reasons. First, as described above, monkeys are poor at recognizing the identity of conspecifics from their faces only. Moreover, even if the animals tested in experiments are usually raised in captivity, there is no reason to expect that they would be better at identity recognition of *human* faces (as used in most of these studies). Second, when they succeed at learning to recognize a modest number of facial identities in photographs, monkeys rely on different cues and qualitatively different processes compared with humans, most likely using a part-based analysis of low-level visual characteristics rather than extracting a holistic, fine-grained representation to individuate the face at a glance. Third, there are major relevant differences at the neural level between humans and macaques (Figure 21.4), who last shared a common ancestor 25–30 million years ago (Kaas, 2019). For a start, the human brain contains about 13.5 times more neurons than the macaque brain (Herculano-Houzel, 2016) and this difference is not homogenous across regions. For instance, the primary visual cortex is relatively much larger in macaques than humans – 10 percent of the total cortical surface versus 3 percent, respectively (Van Essen, 2005). In contrast, the temporal lobe is much smaller and thinner in macaques than humans, even accounting for body size differences (Bryant & Preuss, 2018; Rilling & Seligman, 2002).

Beyond (relative) size, they are also major differences in gyrification between the brains of macaques and humans (Zilles et al., 2013). In humans, the VOTC is divided by two major sulci: the collateral sulcus (CS) more medially, and the occipito-temporal sulcus (OTS) more laterally, these two sulci defining the borders of *the fusiform gyrus*, the most important structure for human FIR, as noted above. In contrast, monkeys possess only one main ventral sulcus, labelled as the OTS, therefore lacking a fusiform gyrus – among other substantial neuroanatomical differences with the human VOTC (Figure 21.4A–B).

Given their lack of the hominoid-specific fusiform structure (Weiner & Zilles, 2016), it is not surprising that macaque monkeys do not hold a ventral face-selective cortical network (Bryant & Preuss, 2018). Rather, their face-selective regions are found more dorsally, mainly in the STS, where face-selective neurons have been recorded in this species since the earliest studies. Critically, bilateral removal of the entire STS in macaque monkeys has no or little effect on their learned ability to discriminate faces of conspecifics for their identity, but disrupts their sensitivity to gaze orientation (Heywood & Cowey, 1992). This latter observation supports the view that these face-selective STS regions in the monkey brain are involved in recognizing dynamic aspects of faces, such as eye gaze direction, head orientation, and facial expression, rather than identity (Allison et al., 2000; Bernstein & Yovel, 2015; Perrett et al., 1985; Taubert et al., 2020).

Finally, in humans, face-selective cortical regions are typically larger in size and level of activation in the right as compared to the left hemisphere (Rossion et al., 2012; Sergent et al., 1992). While the factors driving this right hemispheric lateralization remain largely unknown (Rossion & Lochy, 2022), regions of the right but not the left hemisphere appear critical for FIR (Cohen et al., 2019; Jonas & Rossion, 2021; Sergent & Signoret, 1992). This contrasts with the strictly bilateral cortical face network observed in monkeys (Tsao et al., 2008).

Overall, these anatomo-functional differences between species, together with the lack of human-like ability at FIR in macaque monkeys as reviewed in the previous section, support the exciting view that the ventral face-selective recognition system – which is mainly found all along the lateral fusiform gyrus and adjacent sulci (Figure 21.4) – emerged relatively recently in evolution, and may even have developed specifically in the human species.[3]

The key role of reentrant semantically driven identity representations

If the monkey brain offers only an inadequate model of human FIR, how can we clarify the neural mechanisms of this key function in our species? One answer is technical: neuroscientists should focus in the coming years on electrophysiological recording of neuronal populations and single neurons in the *human* cortical face-network. Human recordings of single neurons have been performed for decades with microelectrodes in patients with epilepsy refractory to medication (Fried et al., 2014). Yet this approach remains very rare and technically challenging. Moreover, for technical and clinical reasons, these studies focus on specific brain structures such as the medial temporal lobe (e.g., hippocampus, amygdala), where so-called concept cells (i.e., neurons responding to multimodal inputs defining a famous person, scene, or object) have been reported (Quian-Quiroga, 2020; Quian-Quiroga et al., 2005). Notwithstanding these limitations, recent studies have started to explore multi- and single-unit spiking activity in face-selective human VOTC regions (Khuvis et al., 2021), providing hope to make progress in our understanding of the neural basis of human FIR in the years to come. However, perhaps more fundamentally, another answer to the question above is conceptual: to make sense of future findings at this level, and integrate them with the large corpus of data collected at multiple levels of brain organization with various techniques of cognitive neuroscience, the prevalent theoretical framework in human face recognition research needs to be fundamentally reevaluated. This is where a neurosemiotic perspective, as advocated in this book and briefly discussed below, can be inspiring.

Human face recognition is generally considered a problem of vision, according to which we need to resolve how the analysis of physical features of a face stimulus impinging on the eye's retina leads to a hierarchical succession of visual representations in the cortex until an invariant visual representation of facial identity is extracted (DiCarlo et al., 2012; Freiwald, 2020; Hesse & Tsao, 2020). It is only *after* this invariant unimodal (visual) representation is built from objective features of the stimulus that it would be associated with semantic knowledge about that face (Bruce & Young, 1986; Jiang et al., 2006; Duchaine & Yovel, 2015). In fact, despite the use of various (and often confusing) terminologies, this dissociation goes well beyond face (identity) recognition and reflects a general conceptual gap that is made in human cognitive (neuro)science between "(visual) perception" and "(semantic) memory," or between "perceptual/visual categorization" (i.e., categorization thought to be based on the decoding of objective physical attributes of stimuli) and "conceptual/semantic categorization" (i.e., abstract categorizations based on our knowledge of the world).

Usually originating from a tradition of research that started with early stages of visual processing (Hubel & Wiesel, 1962), and the first influential computational approach of vision of Marr (1982), monkey neurophysiologists have adopted this hierarchical view of visual recognition, including face recognition, to the core. These researchers consider the visual recognition system as being essentially hierarchical, with increasingly complex visual representations progressively built from the computational analysis and (linear or nonlinear) combination of objective physical features of stimuli (DiCarlo et al., 2012; Freiwald, 2020; Hesse & Tsao, 2020). Since representations of facial identities are thought to be computationally extracted, independently of any previous encounter with these faces, and are based on physically objective features, the inputs (i.e., face images) can even be well-reconstructed from the pattern of neural activity at a specific stage in the system (Chang & Tsao, 2017). According to this hierarchical computational view, memory representations of familiar faces located in more anterior regions of the system, such as the anterior temporal lobe or the parahippocampal gyrus (Landi & Freiwald, 2017), are contacted only following computational extraction of invariant visual representations. The goal of this research program, based on this standard framework, is therefore to elucidate the nature (content) of visual representations at each of the stages of the system and the computations applied to these representations.

However, in our species at least, FIR goes well beyond vision and fully incorporates semantics, including verbal representations. Take the face stimuli of Figure 21.1, for instance: in many instances, physical similarity is much higher between one view of a face identity (George Clooney's face) and another (unfamiliar or familiar) face identity than between two views of George Clooney's face. This is even clearer when comparing veridical facial identities to their caricatures or their lookalikes. In this context, how could the visual recognition system build an invariant representation of all views of a face identity (e.g., George Clooney), independently of knowledge (semantic memory) about this person? And while there is probably no lookalike that we would seriously confuse with Donald Trump's inimitable face, how do we readily recognize him as a Parisian trash bin?[4] These examples serve to illustrate that conceptualizing FIR as a problem of vision only, with (semantic) memory representations intervening afterwards, is not tenable. This standard position in human face recognition research, according to which a series of perceptual stages precedes access to semantic knowledge, leads to fundamental incoherencies, such as the rejection of unfamiliar FIR to non-expert low-level visual processes (Megreya & Burton, 2006; Young & Burton, 2018), despite overwhelming evidence that such faces are processed holistically in high-level human visual cortical areas (Jacques et al., 2020; see Rossion, 2018b). Moreover, this standard position of a dissociation between perceptual processing stages and semantic memory makes it virtually impossible to understand how unfamiliar faces can (rapidly) become familiar.

To resolve these issues, an alternative view is proposed here according to which semantic representations – that is, the source of information we possess that enables us to bring meaning to words, objects, and all other nonverbal stimuli such as smell or sound (Lambon-Ralph et al., 2010) – in the ventral anterior temporal lobe developmentally *precede* and *build* unimodal (i.e., visual) face-selective representations of identity in the posterior VOTC, *constraining* these representations throughout lifetime via reentrant neuronal connections. That is, one learns that different views of a face belong to the same identity not only because they often occur in close temporal succession (Wallis & Bülthoff, 2001), but because they are associated with the same context, affects, auditory, and even olfactory inputs (Rekow et al., 2021), as well as abstract verbal cues – for example, names and all kinds of semantic associations (Dixon et al., 1998; Eick et al., 2020; Schwartz & Yovel, 2016). This is how two completely different views of an unknown George Clooney can become associated in our brain: thanks to their non-visual semantic commonalities, they both initially lead to the firing of a largely overlapping population of neurons in the ventral anterior temporal cortex because we encounter them in common multimodal contexts. And, contrary to the standard model of human face recognition, this multimodal person representation is not contacted *after* the extraction of view-invariant visual representations: to the contrary, it constitutes the earliest – in terms of developmental origin and learning – and (only fully) invariant representations of "facial" identity. In humans, these multimodal person representations are located in anterior regions of the VOTC – that is, the ventral anterior temporal lobe (Collins et al., 2016) – being heavily constrained by language-related inputs.

Through re-entry – that is, bidirectional rapid exchange of inputs along reciprocal axonal fibers (Edelman & Gally, 2013) – these selective populations of neurons synchronize their activity with populations of neurons located more posteriorly in the VOTC, down to non-retinotopic visual regions of the inferior occipital gyrus, posterior and midfusiform gyrus (Figure 21.5). If multimodal recognition in ventral anterior temporal regions is successful, this leads to strengthening of connections with selective populations of neurons in posterior VOTC regions – a process occurring through Hebbian associative learning mediated by the hippocampus (Squire et al., 2015). Since these posterior VOTC regions receive concurrent direct inputs from (retinotopic) low-level visual areas, there is a progressive selection of their neuronal populations receiving the patterns of low-level visual inputs associated with successful recognition. Through this associative process, in these posterior VOTC regions, facial identities become progressively distinguishable based on low-level visual inputs only. Thus, unimodal (i.e., visual) cortical memories (Fuster, 2009) are built "inside-out," from semantics to vision. In a mature system, these cortical visual memories are essential to immediately distinguish novel faces from one another, even in the absence of semantic associations, and facilitate learning – that is, transforming unfamiliar faces into familiar ones.

Importantly, this dynamic recurrent process leads to *distributed* cortical memories (of facial identities) all along the human VOTC (Figure 21.5). Rather than fixed hierarchical computational stages with an invariant, purely visual representation extracted from the input and subsequently associated with semantic memory, there is an antero-posterior *gradient* of selected semantic-visual feature representation (Figure 21.5). Hence, even "visual" posterior VOTC face-selective regions can be modulated in their early activity by multimodal inputs linked to faces and facial identities, such as human voices (Blank et al., 2015), identity-related names (Ambrus et al., 2019), semantic associations (e.g., "job title" – Eick et al., 2020) or even face-related words (García et al., 2020). Yet, in a fully

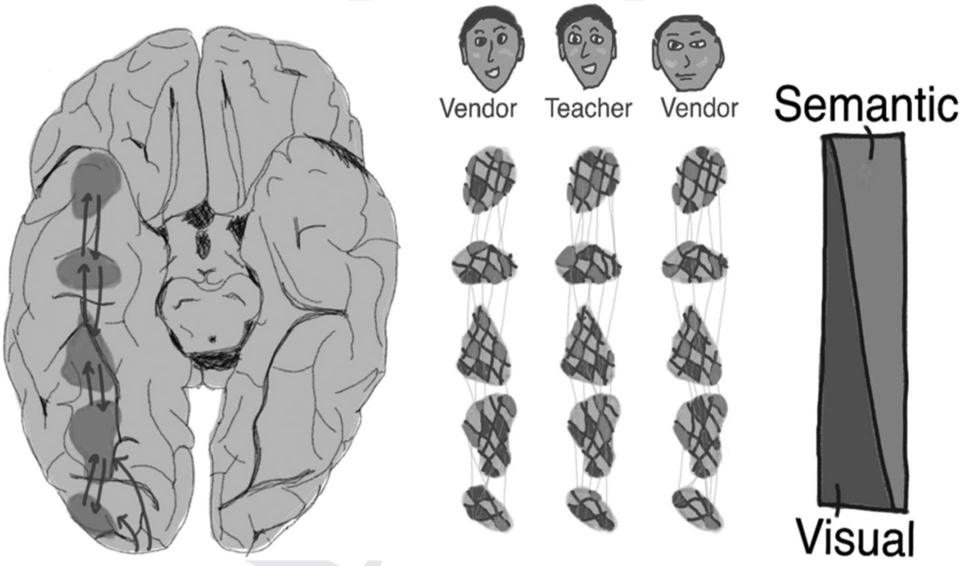

Figure 21.5 Face identity recognition in the human ventral occipito-temporal cortex

Different face-selective populations of neurons ("clusters") are linked by reentrant connections. Through a gradient of sensitivity from purely visual to semantic features along the VOTC, different facial identities could be represented by partly overlapping groups of neurons within these clusters, providing increased invariance across physical variations in the antero-posterior direction. Note that while five clusters of roughly similar sizes are represented here, their number and spatial extent vary across individual brains (Figure 21.4C). See e-book for a full-color version of this figure.

developed neurotypical adult brain, their main source of inputs to distinguish identities comes from low-level visual regions. This is the reason why, in a mature system, selective destruction of these posterior VOTC face-selective regions – such as the "FFA" in the right hemisphere – may cause prosopagnosia, a modality-specific memory impairment limited to face identity (Rossion, 2018a).

Finally, another implication of this visuo-semantic inverse gradient is that, although there are general anatomical constraints for these selective populations of neurons to end up roughly in the same cytoarchitectonic VOTC regions – for example, the lateral rather than the medial fusiform gyrus (Grill-Spector et al., 2017) – they also vary tremendously in terms of variability, size, spatial distribution, and number across individual brains (Figure 21.4C) (Gao et al., 2022), reflecting specific cultural and idiosyncratic visuo-semantic experiences with faces. Consequently, any scientific attempt to define general computational/representational stages at each level of a well-defined hierarchy of brain areas across all individual brains is bound to fail.

While a full elaboration of this original view goes beyond the scope of this chapter, if proved correct it would firmly place human face (identity) recognition beyond the realm of vision, as part of the larger field of neurosemiotics.

Conclusions: Towards a neurosemiotics of human face recognition?

A fundamental function of the central nervous system is to recognize stimuli of the environment in order to generate adaptive behavior. This function is not achieved by a series of computations applied to well-defined distinct representational stages but "merely" by associating specific populations of neurons to specific stimuli and being able to reproduce these selective responses across variable instances of the stimuli. At this very coarse level of description, there is no fundamental difference between the recognition of the identity of conspecifics from the face, FIR, across a number of animal species – for example, Polistes fuscatus wasps, sheep, macaque monkeys and humans. In humans, however, FIR requires a very fine-grained level of visual discrimination among many exemplars and, especially immense generalization abilities that go well beyond objective physical features. Thanks to extensive experience during development and the social pressure to rapidly and accurately recognize the identity of a large and variable number of face identities, this leads to an unparaleled level of expertise at FIR in neurotypical human adults, characterized by high accuracy levels, speed, automaticity and holistic recognition. This FIR expertise is supported by a distributed category-selective right-lateralized network in the VOTC, a network that is not found in macaque monkeys and that, across individual human brains, presents with a large amount of interindividual variability in size, localization, and number of cortical nodes (Gao et al., 2022). The VOTC face network originates as a semantic selective neural response in the ventral anterior temporal lobe, gradually generating and driving visual face-selectivity in posterior regions through re-entrant connections to support human-specific expertise at FIR (Figure 21.5).

According to this view, face (identity) recognition, even for unfamiliar faces, necessarily concerns meaning in the nervous system rather than a reactive information processing decoding, making it a semiotic process (Queiroz & El-Hani, 2006). That is, what determines the diagnosticity of low-level visual inputs (i.e., their "ability" to lead to successful recognition of a face or of a specific face identity) is not only their objective content (i.e., features to "decode"), but how these inputs are (and have been) successfully associated with semantic knowledge (in the past). Since it is not merely an object to decode visually, but a sign that is always interpreted based on past and present successful associations in the nervous system, this perspective puts the human face firmly in the field of neurosemiotics.

Acknowledgements

I thank Talia Retter for editing the manuscript and generating Figure 21.5, Jacques Jonas for contributing to Figure 21.4 and in-depth discussion of the comparative neuroanatomy of the human and

macaque brain, Marie-Alphée Laurent and Benoit Cottereau for the macaque monkey activation maps in Figure 21.4, and Adolfo García and Agustin Ibáñez for their helpful comments on a previous version of this chapter.

Notes

1 www.theguardian.com/uk-news/2018/may/15/uk-police-use-of-facial-recognition-technology-failure

2 National Institute of Standards and Technology Interagency or Internal Report 8311.

3 Despite these obvious differences, a number of prominent scientists in the field consider that the monkey brain holds the exact same number of face-selective regions as the human brain (Tsao et al., 2008; Weiner & Grill-Spector, 2015) and that the STS cortical face network in the monkey brain should be divided into two parts: (1) a dorsal part, homologous to the face-selective STS network in humans; and (2) a ventral part, corresponding to the human face-selective VOTC network (Freiwald et al., 2016; Tsao et al., 2008; Weiner & Grill-Spector, 2015; Yovel & Freiwald, 2013) (Figure 21.4). According to this view, the putative ventral part of the monkey STS system would have been "pushed" onto the ventral cortical surface over the course of human evolution by the expansion of regions engaged in language and social cognition, in particular the temporal parietal junction. However, the hypothesis of a one-to-one homology of regions across the two species and of a ventral displacement of lower STS clusters through human evolution is far-fetched and contradicted by considerable anatomo-functional evidence (Gao et al., 2022; Rossion & Taubert, 2019), falling flat in comparison with the wealth of converging evidence supporting the view that the face-selective VOTC network is indeed specific to the human lineage.

4 www.boredpanda.com/funny-donald-trump-bin-photoshop-battle/?utm_source=google&utm_medium=organic&utm_campaign=organic

References

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: Role of the STS region. *Trends in Cognitive Sciences*, *4*, 267–278. doi: 10.1016/s1364-6613(00)01501-1

Ambrus, G.G., Amado, C., Krohn, L., & Kovacs, G. (2019). TMS of the occipital face area modulates cross-domain identity priming. *Brain Structure and Function*, *224*, 149–157. doi: 10.1007/s00429-018-1768-0

Barton, J. J., Davies-Thompson, J., & Corrow, S. L. (2021). Prosopagnosia and disorders of face processing. In *Handbook of Clinical Neurology*, *Vol. 178* (pp. 175–193). New York: Elsevier.

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

Bernstein, M. & Yovel, G. (2015). Two neural pathways of face processing: A critical evaluation of current models. *Neuroscience & Biobehavioral Reviews*, *55*, 536–546. doi: 10.1016/j.neubiorev.2015.06.010

Blank, H., Kiebel, S. J., & von Kriegstein, K. (2015). How the human brain exchanges information across sensory modalities to recognize other people. *Human Brain Mapping*, *36*, 324–39. doi: 10.1002/hbm.22631

Bodamer, J. (1947). Die Prosopagnosie. *Archiv fur Psychiatrie und Zeitschrift fur Neurologie*, *179*, 6–54. doi: 10.1007/BF00352849

Bruce, C. (1982). Face recognition by monkeys: Absence of an inversion effect. *Neuropsychologia*, *20*, 515–521. doi: 10.1016/0028-3932(82)90025-2

Bruce, V., Henderson, Z., Newman, C., & Burton, A. M. (2001). Matching identities of familiar and unfamiliar faces caught on CCTV images. *Journal of Experimental Psychology: Applied*, *7*, 207–218. doi: 10.1037//1076-898X.7.3.207

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, 77, 305–327. doi: 10.1111/j.2044-8295.1986.tb02199.x

Bruce, V., & Young, A. (1998). *In the Eye of the Beholder: The Science of Face Perception*. New York: Oxford University Press.

Bryant, K. & Preuss, T. (2018). A comparative perspective on the human temporal lobe. In E. Bruner, N. Ogihara, & H. Tanabe (Eds.), *Digital Endocasts: From Skulls to Brains* (pp. 239–258). New York: Springer.

Burton, A. M., Kramer, R. S., Ritchie, K. L., & Jenkins, R. (2016). Identity from variation: Representations of faces derived from multiple instances. *Cognitive Science*, *40*, 202–223. doi: 10.1111/cogs.12231

Calder, A. J., Rhodes, G., Johnson, M. H., & Haxby, J. V. (Eds). (2011). *The Oxford Handbook of Face Perception*. Oxford: Oxford University Press.

Carey, S. (1992). Becoming a face expert. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *335*, 95–103. doi: 10.1098/rstb.1992.0012

Cavalli-Sforza, L. L. & Cavalli-Sforza, F. (1995). *The Great Human Diasporas: The History of Diversity and Evolution*. Massachusetts: Addison-Wesley.

Chang, L. & Tsao, D. Y. (2017). The code for facial identity in the primate brain. *Cell*, *169*, 1013–1028. doi: 10.1016/j.cell.2017.05.011

Cohen, A. L., Soussand, L., Corrow, S. L., Martinaud, O., Barton, J. J. S., & Fox, M. D. (2019). Looking beyond the face area: Lesion network mapping of prosopagnosia. *Brain*, *142*, 3975–3990. doi: 10.1093/brain/awz332

Collins, J. A., Koski, J. E., & Olson, I. R. (2016). More than meets the eye: The merging of perceptual and conceptual knowledge in the anterior temporal face area. *Frontiers in Human Neuroscience*, *10*, 189. doi: 10.3389/fnhum.2016.00189

Crookes, K. & McKone, E. (2009). Early maturity of face recognition: No childhood development of holistic processing, novel face encoding, or face-space. *Cognition*, *111*, 219–247. doi: 10.1016/j.cognition.2009.02.004

Davies-Thompson, J., Gouws, A., & Andrews, T. J. (2009). An image-dependent representation of familiar and unfamiliar faces in the human ventral stream. *Neuropsychologia*, *47*, 1627–1635. doi: 10.1016/j.neuropsychologia.2009.01.017

Desimone, R. (1991). Face-selective cells in the temporal cortex of monkeys. *Journal of Cognitive Neuroscience*, *3*, 1–8. doi: 10.1162/jocn.1991.3.1.1

Diamond, R. & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General*, *115*, 107–117. doi: 10.1037//0096-3445.115.2.107

DiCarlo, J. J., Zoccolan, D., & Rust, N. C. (2012). How does the brain solve visual object recognition? *Neuron*, *73*, 415–34. doi: 10.1016/j.neuron.2012.01.010

Dixon, M. J., Bub, D. N, & Arguin, M. (1998). Semantic and visual determinants of face recognition in a prosopagnosic patient. *Journal of Cognitive Neuroscience*, *10*, 362–376. doi: 10.1162/089892998562799

Duchaine, B. & Yovel, G. (2015). A revised neural framework for face processing. *Annual Review of Vision Science, 1*, 293–416. doi: 10.1146/annurev-vision-082114-035518

Edelman, G. M. (2004). Biochemistry and the sciences of recognition. *Journal of Biological Chemistry*, *279*, 7361–7369. doi: 10.1074/jbc.X400001200

Edelman, G. M. & Gally, J. A. (2013). Reentry: A key mechanism for integration of brain function. *Frontiers in Integrative Neuroscience*, 7, 63. doi: 10.3389/fnint.2013.00063

Eick, C. M., Kovacs, G., Rostalski, S. M., Rohrig, L., & Ambrus, G. G. (2020). The occipital face area is causally involved in identity-related visual-semantic associations. *Brain Structure and Function*, *225*, 1483–1493. doi: 10.1007/s00429-020-02068-9

Emery, N. J., Lorincz, E. N., Perrett, D. I., Oram, M. W., & Baker, C. I. (1997). Gaze following and joint attention in rhesus monkeys (Macaca mulatta). *Journal of Comparative Psychology*, *111*, 286–93. doi: 10.1037/0735-7036.111.3.286

Farkas, L. G. (1994). *Anthropometry of the Head and Face*. New York: Raven Press.

Felleman, D. J. & van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, *1*, 11–47. doi: 10.1093/cercor/1.1.1-a

Freiwald, W. A. (2020). The neural mechanisms of face processing: Cells, areas, networks, and models. *Current Opinion in Neurobiology*, *60*, 184–191. doi: 10.1016/j.conb.2019.12.007

Freiwald, W., Duchaine, B., & Yovel, G. (2016). Face processing systems: From neurons to real-world social perception. *Annual Review of Neuroscience*, *39*, 325–346. doi: 10.1146/annurev-neuro-070815-013934

Fried, I., Rutishauser, U., Cerf, M., & Kreiman, G. (2014). In I. Fried, U. Rutishauser, M. Cerf, & G. Kreiman (Eds.), *Single Neuron Studies of the Human Brain: Probing Cognition*. Cambridge, MA: MIT Press.

Fuster, J. M. (2009). Cortex and memory: Emergence of a new paradigm. *Journal of Cognitive Neuroscience*, *21*, 2047–2072. doi: 10.1162/jocn.2009.21280

Gainotti, G. (2013). Is the right anterior temporal variant of prosopagnosia a form of "associative prosopagnosia" or a form of "multimodal person recognition disorder"? *Neuropsychology Review*, *23*, 99–110. doi: 10.1007/s11065-013-9232-7

Gao, X., Gentile, F., & Rossion, B. (2018). Fast periodic stimulation (FPS): A highly effective approach in fMRI brain mapping. *Brain Structure and Function*, *223*, 2433–2454. doi: 10.1007/s00429-018-1630-4

Gao, X., Wen, M., Sun, M., & Rossion, B. (2022). A genuine interindividual variability in number and anatomical localization of face-selective regions in the human brain. *Cerebral Cortex*, bhab519. doi: 10.1093/cercor/bhab519

García, A. M., Hesse, E., Birba, A., Adolfi, F., Mikulan, E., Caro, M. M., Petroni, A., Bekinschtein, T. A., García, M. C., Silva, W., Ciraolo, C., Vaucheret, E., Sedeño, L., & Ibáñez, A. (2020). Time to face language: Embodied mechanisms underpin the inception of face-related meanings in the human brain. *Cerebral Cortex*, *30*, 6051–6058. doi: 10.1093/cercor/bhaa178

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

Gauthier, I., Tarr, M. J., Moylan, J., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). The FFA is part of a network that processes faces on an individual level. *Journal of Cognitive Neuroscience*, *12*, 495–504. doi: 10.1162/089892900562165

Gazzaniga, M. S. (2009). *Human: The Science Behind What Makes Us Unique*. New York: Harper Perennial.

Golarai, G., Liberman, A., & Grill-Spector, K. (2017). Experience shapes the development of neural substrates of face processing in human ventral temporal cortex. *Cerebral Cortex*, *27*, 1229–1244. doi: 10.1093/cercor/bhv314

Griffin, J. W. (2020). Quantifying the face inversion effect in nonhuman primates: A phylogenetic meta-analysis. *Animal Cognition*, *23*, 237–249. doi: 10.1007/s10071-019-01340-8

Grill-Spector, K., & Malach, R. (2001). FMR-adaptation: A tool for studying the functional properties of human cortical neurons. *Acta Psychologica*, *107*, 293–321. doi: 10.1016/s0001-6918(01)00019-1

Grill-Spector, K., Weiner, K. S., Kay, K., & Gomez, J. (2017). The functional neuroanatomy of human face perception. *Annual Review of Vision Science*, *3*, 167–196. doi: 10.1146/annurev-vision-102016-061214

Gross, C. G., Rocha-Miranda, C. E., & Bender, D. B. (1972). Visual properties of neurons in the inferotemporal cortex of the macaque. *Journal of Neurophysiology*, *35*, 96–111. doi: 10.1152/jn.1972.35.1.96

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

Herculano-Houzel, S. (2016). *The Human Advantage: A New Understanding of How Our Brain Became Remarkable*. Cambridge, MA: MIT Press.

Hesse, J. K. & Tsao, D.Y. (2020). The macaque face patch system: A turtle's underbelly for the brain. *Nature Reviews Neuroscience*, *21*, 695–716. doi: 10.1038/s41583-020-00393-w

Heywood, C. A. & Cowey, A. (1992). The role of the "face-cell" area in the discrimination and recognition of faces by monkeys. *Philosophical Transactions of the Royal Society of London B*, *335*, 31–37. doi: 10.1098/rstb.1992.0004

Hills, P. J. & Lewis, M. B. (2018). The development of face expertise: Evidence for a qualitative change in processing. *Cognitive Development*, *48*, 1–18. doi: 10.1016/j.cogdev.2018.05.003

Hsiao, J. H. & Cottrell, G. W. (2008). Two fixations suffice in face recognition. *Psychological Science*, *19*, 998–1006. doi: 10.1111/j.1467-9280.2008.02191.x

Hubel, D. H. & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat visual cortex. *The Journal of Physiology*, *160*, 106–154. doi: 10.1113/jphysiol.1962.sp006837

Hutmacher, F. (2019). Why is there so much more research on vision than on any other sensory modality? *Frontiers in Psychology*, *10*, 2246. doi: 10.3389/fpsyg.2019.02246

Jacques, C., Rossion, B., Volfart, A., Brissart, H., Colnat-Coulbois, S., Maillard, L., Jonas, J. (2020). The neural basis of rapid unfamiliar face individuation with human intracerebral recordings. *NeuroImage*, 221, 117174. doi: 10.1016/j.neuroimage.2020.117174

Jenkins, R., Dowsett, A. J., & Burton, A. M. (2018). How many faces do people know? *Proceedings of the Royal Society B, Biological Sciences*, *285*(1888). doi: 10.1098/rspb.2018.1319

Jiang, X., Rosen, E., Zeffiro, T., Vanmeter, J., Blanz V., & Riesenhuber, M. (2006). Evaluation of a shape-based model of human face discrimination using fMRI and behavioral techniques. *Neuron*, *50*, 159–172. doi: 10.1016/j.neuron.2006.03.012

Jonas, J., Jacques, C., Liu-Shuang, J., Brissart, H., Colnat-Coulbois, S., Maillard, L., & Rossion, B. (2016). A face-selective ventral occipito-temporal map of the human brain with intracerebral potentials. *Proceedings of the National Academy of Sciences of the United States of America*, *113*, E4088–E4097. doi: 10.1073/pnas.1522033113

Jonas, J., & Rossion, B. (2021). Intracerebral electrical stimulation to understand the neural basis of human face identity recognition. *European Journal of Neuroscience*, 54(1), 4197–4211. doi: 10.1111/ejn.15235

Kaas, J. H. (2019). The origin and evolution of neocortex: From early mammals to modern humans. *Progress in Brain Research*, *250*, 61–81. doi: 10.1016/bs.pbr.2019.03.017

Kanwisher, N., McDermott, J., & Chun, M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, *17*, 4302–4311. doi: 10.1523/JNEUROSCI.17-11-04302.1997

Kelly, D. J., Liu, S., Ge, L., Quinn, P. C., Salater, A. M., Lee, K., Liu, Q., & Pascalis, O. (2007). Cross-race preferences for same-race faces extend beyond the African versus Caucasian contrast in three-month-old infants. *Infancy*, *11*, 87–95. doi: 10.1080/15250000709336871

Khuvis, S., Yeagle, E. M., Norman, Y., Grossman, S., Malach, R., & Mehta, A. (2021). Face-selective units in human ventral temporal cortex reactivate during free recall. *Journal of Neuroscience*, *41*, 3386–3399. doi: 10.1523/JNEUROSCI.2918-19.2020

Lambon-Ralph M. A., Sage, K., Jones, R. W., & Mayberry, E. J. (2010). Coherent concepts are computed in the anterior temporal lobes. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(6), 2717–2722. doi: 10.1073/pnas.0907307107

Landi, S. M., & Freiwald, W. A. (2017). Two areas for familiar face recognition in the primate brain. *Science*, *357*(6351), 591–595. doi: 10.1126/science.aan1139

Le Grand, R., Mondloch, C., Maurer, D., & Brent, H. (2004). Impairment in holistic face processing following early visual deprivation. *Psychological Science*, *15*, 762–768. doi: 10.1111/j.0956-7976.2004.00753.x

Leopold, D. A., Bondar, I. V., & Giese, M.A. (2006). Norm-based face encoding by single neurons in monkey infero-temporal cortex. *Nature, 442*, 572–575. doi: 10.1038/nature04951

Logothetis, N. K. (2008). What we can do and what we cannot do with fMRI. *Nature*, *453*, 869–878. doi: 10.1038/nature06976

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, *87*, 252–271. doi: 10.1037/0033-295X.87.3.252

Marr, D. (1982). *Vision: A Computational Investigation into Human Representation and Processing of Visual Information*. San Francisco, CA: Freeman.

McKone, E., Kanwisher, N., & Duchaine, B. C. (2007). Can generic expertise explain special processing for faces? *Trends in Cognitive Sciences*, *11*, 8–15. doi: 10.1016/j.tics.2006.11.002

McNeill, D. (2000). *The Face: A Natural History*. Boston: Little, Brown.

Meadows, J. C. (1974). The anatomical basis of prosopagnosia. *Journal of Neurology, Neurosurgery, and Psychiatry*, *37*, 489–501. doi: 10.1136/jnnp.37.5.489

Megreya, A. M. & Burton, A. M. (2006). Unfamiliar faces are not faces: Evidence from a matching task. *Memory & Cognition*, *34*, 865–876. doi: 10.3758/BF03193433

Micheletta, J., Whitehouse, J., Parr, L. A., Marshman, P., Engelhardt, A., & Waller, B. M. (2015). Familiar and unfamiliar face recognition in crested macaques (Macaca nigra). *Royal Society Open Science, 2*, 150109. doi: 10.1098/rsos.150109

Murphy, G. L. (2002). *The Big Book of Concepts*. Cambridge, MA: MIT Press.

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. doi: 10.1073/pnas.1102433108

Palmeri, T. J., & Gauthier, I. (2004). Visual object understanding. *Nature Reviews Neuroscience*, *5*, 291–303. doi: 10.1038/nrn1364

Parr, L. A., Heintz, M., & Pradhan, G. (2008). Rhesus monkeys (Macaca mulatta) lack expertise in face processing. *Journal of Comparative Psychology*, *122*, 390–402. doi: 10.1037/0735-7036.122.4.390

Pascalis, O., de Haan, M., & Nelson, C. A. (2002). Is face processing species-specific during the first year of life? *Science*, *296*(5571), 1321–1323. doi: 10.1126/science.1070223

Peirce, J. W., Leigh, A. E., & Kendrick, K. M. (2000). Configurational coding, familiarity and the right hemisphere advantage for face recognition in sheep. *Neuropsychologia*, *38*, 475–483. doi: 10.1016/s0028-3932(99)00088-3

Perrett, D. I., Rolls E. T., & Caan W. (1982). Visual neurones responsive to faces in the monkey temporal cortex. *Experimental Brain Research*, *47*, 329–342. doi: 10.1007/BF00239352

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 B*, 223, 293–317. doi: 10.1098/rspb.1985.0003

Pokorny, J. J., & de Waal, F. B. (2009). Face recognition in capuchin monkeys (Cebus apella). *Journal of Comparative Psychology*, *123*, 151–160. doi: 10.1037/a0014073

Queiroz, J., & El-Hani, C. (2006). Towards a multi-level approach to the emergence of meaning processes in living systems. *Acta Biotheoretica*, *54*, 179–206. doi: 10.1007/s10441-006-8177-0

Quian-Quiroga, R. (2020). No pattern separation in the human hippocampus. *Trends in Cognitive Sciences*, *24*, 994–1007. doi: 10.1016/j.tics.2020.09.012

Quian-Quiroga, R., Reddy, L., Kreiman, G., Koch, C, & Fried, I. (2005). Invariant visual representation by single neurons in the human brain. *Nature*, *435*, 1102–1107. doi: 10.1038/nature03687

Ramon, M., Busigny, T., & Rossion, B. (2010). Impaired holistic processing of unfamiliar individual faces in acquired prosopagnosia. *Neuropsychologia*, *48*, 933–944. doi: 10.1016/j.neuropsychologia.2009.11.014

Rekow, D., Baudouin, J.Y., Poncet, F., Damon, F., Durand, K., Schaal, B., Rossion, B., & Leleu, A. (2021). Smells like real faces: Odor-driven face pareidolia in the infant brain. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(21), e2014979118. doi: 10.1073/pnas.2014979118

Rilling, J. K. & Seligman, R. A. (2002): A quantitative morphometric comparative analysis of the primate temporal lobe. *Journal of Human Evolution*, *42*, 505–533. doi: 10.1006/jhev.2001.0537

Rolls, E. T. (1992). Neurophysiological mechanisms underlying face processing within and beyond the temporal cortical area. *Philosophical Transactions of the Royal Society of London B*, *335*, 11–21. doi: 10.1098/rstb.1992.0002

Rossion, B. (2008). Picture-plane inversion leads to qualitative changes of face perception. *Acta Psychologica*, *128*, 274–289. doi: 10.1016/j.actpsy.2008.02.003

Rossion, B. (2013). The composite face illusion: A window to our understanding of holistic face perception. *Visual Cognition*, *121*, 139–253. doi: 10.1080/13506285.2013.772929

Rossion, B. (2014). Understanding face perception by means of prosopagnosia and neuroimaging. *Frontiers in Bioscience*, *6*, 258–307. doi: 10.2741/E706

Rossion, B. (2018a). Humans are visual experts at unfamiliar face recognition. *Trends in Cognitive Sciences*, *22*, 471–472. doi: 10.1016/j.tics.2018.03.002

Rossion, B. (2018b). Damasio's error – prosopagnosia with intact within-category object recognition. *Journal of Neuropsychology*, *12*, 357–388. doi: 10.1111/jnp.12162

Rossion, B. (2018c). Prosopdysgnosia? What could it tell us about the neural organization of face and object recognition? *Cognitive Neuropsychology*, *35*, 98–101. doi: 10.1080/02643294.2017.1414778

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. doi: 10.1016/j.bandc.2012.01.001

Rossion, B., Jacques, C., & Jonas, J. (2018). Mapping face categorization in the human ventral occipito-temporal cortex with direct neural intracranial recordings. *Annals of the New York Academy of Science*, *1426*, 5–24. doi: 10.1111/nyas.13596

Rossion, B., & Lochy, A. (2022). Is human face recognition lateralized to the right hemisphere due to neural competition with left lateralized visual word recognition? A critical review. *Brain Structure and Function*, 227(2), 599–629. doi: 10.1007/s00429-021-02370-0

Rossion, B., & Retter, T. L. (2020). Face perception. In D. Poeppel, M.S. Gazzaniga & G. Mangun (Eds). *The Cognitive Neurosciences* (6th edn, pp. 129–139). Cambridge, MA: MIT Press.

Rossion, B., & Taubert, J. (2017). Commentary: The code for facial identity in the primate brain. *Frontiers in Human Neuroscience*, *11*, 550. doi: 10.3389/fnhum.2017.00550

Rossion, B., & Taubert, J. (2019). What can we learn about human individual face recognition from experimental studies in monkeys? *Vision Research*, *157*, 142–158. doi: 10.1016/j.visres.2018.03.012

Ruff, C. (2002). Variation in human body size and shape. *Annual Review of Anthropology*, *31*, 211–232. doi: 10.1146/annurev.anthro.31.040402.085407

Sangrigoli, S. & de Schonen, S. (2004). Effect of visual experience on face processing: A developmental study of inversion and non-native effects. *Developmental Science*, 7, 74–87. doi: 10.1111/j.1467-7687.2004.00324.x

Schwartz, L. & Yovel G. (2016). The roles of perceptual and conceptual information in face recognition. *Journal of Experimental Psychology: General*, *145*, 1493–1511. doi: 10.1037/xge0000220

Sergent, J., Ohta, S., & Macdonald, B. (1992). Functional neuroanatomy of face and object processing: A positron emission tomography study. *Brain*, *115*, 15–36. doi: 10.1093/brain/115.1.15

Sergent, J., & Signoret, J.-L. (1992). Varieties of functional deficits in prosopagnosia. *Cerebral Cortex*, *2*, 375–388. doi: 10.1093/cercor/2.5.375

Sheehan, M. J., & Nachman, M. W. (2014). Morphological and population genomic evidence that human faces have evolved to signal individual identity. *Nature Communications*, *5*(4800), 1–10. doi: 10.1038/ncomms5800

Sheehan, M. J., & Tibbetts, E. A. (2011). Specialized face learning is associated with individual recognition in paper wasps. *Science*, *334*, 1272–1275. doi: 10.1126/science.1211334

Squire, L. R., Genzel L, Wixted, J. T., & Morris, R. G. (2015). Memory consolidation. *Cold Spring Harbor Perspectives in Biology*, 7(8), a021766. doi: 10.1101/cshperspect.a021766

Susilo, T., & Duchaine, B. (2013). Advances in developmental prosopagnosia research. *Current Opinion in Neurobiology*, *23*(3), 423–429. doi: 10.1016/j.conb.2012.12.011

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology*, *46A*, 225–245. doi: 10.1080/14640749308401045

Taubert, J., Aagten-Murphy, D., & Parr, L. A. (2012). A comparative study of face processing using scrambled faces. *Perception*, *41*, 460–473. doi: 10.1068/p7151

Taubert, J., Japee, S., Murphy, A. P., Tardiff, C. T., Koele, E. A., Kumar, S, Leopold, D. A., & Ungerleider, L. G. (2020). Parallel processing of facial expression and head orientation in the macaque brain. *Journal of Neuroscience*, *40*, 8119–8131. doi: 10.1523/JNEUROSCI.0524-20.2020

Taubert, J., Qureshi, A. A., & Parr, L. A. (2012). The composite face effect in chimpanzees (*Pan troglodytes*) and rhesus monkeys (*Macada mulatta*). *Journal of Comparative* Psychology, *126*, 339–46. doi: 10.1037/a0027287

Todorov, A. (2017). *Face Value: The Irresistible Influence of First Impressions*. Princeton, NJ: Princeton University Press.

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. doi: 10.1126/science.1119983

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*, 19514–19519. doi: 10.1073/pnas.0809662105

Van Belle, G., de Graef, P., Verfaillie, K., Busigny, T., & Rossion, B. (2010). Whole not hole: Expert face recognition requires holistic perception. *Neuropsychologia*, *48*, 2609–2620. doi: 10.1016/j.neuropsychologia.2010.04.034

Van Essen, D. C. (2005) Surface-based comparisons of macaque and human cortical organization. In S. Dehaene (Ed.), *From Monkey to Brain: Proceedings of a Fyssen Foundation Colloquium* (3–20). Cambridge, MA: MIT Press.

Wallis, G., & Bülthoff, H. H. (2001). Effects of temporal association on recognition memory. *Proceedings of the National Academy of Sciences of the United States of America*, *98*, 4800–4804. doi: 10.1073/pnas.071028598

Wandell B. A. (2011). The neurobiological basis of seeing words. *Annals of the New York Academy of Sciences*, *1224*(1), 63–80. doi: 10.1111/j.1749-6632.2010.05954.x

Wang, M.Y., & Takeuchi, H. (2017). Individual recognition and the "face inversion effect" in medaka fish (Oryzias latipes). *eLife*, *6*(e24728), 1–18. doi: 10.7554/eLife.24728

Weiner, K. S., & Grill-Spector, K. (2015). The evolution of face processing networks. *Trends in Cognitive Sciences*, *19*, 240–241. doi: 10.1016/j.tics.2015.03.010

Weiner, K. S., & Zilles, K. (2016). The anatomical and functional specialization of the fusiform gyrus. *Neuropsychologia*, *83*, 48–62. doi: 10.1016/j.neuropsychologia.2015.06.033

Wilmer, J. B. (2017). Individual differences in face recognition: A decade of discovery. *Current Directions in Psychological Science, 26*(3), 225–230. doi: 10.1177/0963721417710693

Wilmer, J. B., Germine, L., Chabris, C. F., Chatterjee, G., Williams, M., Loken, E., Nakayama, K., & Duchaine, B. (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

Yan, X., & Rossion, B. (2020). A robust neural familiar face recognition response in a dynamic (periodic) stream of unfamiliar faces. *Cortex*, *132*, 281–295. doi: 10.1016/j.cortex.2020.08.016

Young, A. W., & Burton, M. (2018). Are we face experts? *Trends in Cognitive Sciences*, *22*, 100–110. doi: 10.1016/j.tics.2017.11.007

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configural information in face perception. *Perception*, *16*, 747–759. doi: 10.1068/p160747

Young, M.P., & Yamane, S. (1992). Sparse population coding of faces in the inferotemporal cortex. *Science*, *29*, 1327–1331. doi: 10.1126/science.1598577

Yovel, G., & Freiwald, W. A. (2013). Face recognition systems in monkey and human: Are they the same thing? *F1000Prime Rep*, *5*, 10. doi: 10.12703/P5-10

Zaidi, A. A., Mattern, B. C, Claes, P., McEvoy, B., Hughes, C., & Shriver, M. D. (2017). Investigating the case of human nose shape and climate adaptation. *PLOS Genetics*, (3), e1006616. doi: 10.1371/journal.pgen.1006616

Zilles, K., Palomero-Gallagher, N., & Amunts, K. (2013). Development of cortical folding during evolution and ontogeny. *Trends in Neurosciences*, *36*, 275–284. doi: 10.1016/j.tins.2013.01.006