Chapter 2
Cognitive and Neural Aspects of Face Processing

Faces represent the stimuli we rely on the most for social interaction. They inform us about the identity, mood, gender, age, attractiveness, race and approachability of a person. This is remarkable if we think that all faces share the same composition of internal features (i.e., two eyes above the nose and a mouth) and 3D structure. Thus, faces are unique in terms of the richness of social signals they convey, and the reason why face perception has played a central role for social interaction in a wide range of species for millions of years. Given its importance, face processing has also become one of the most prominent areas of research in cognitive science of the last 50 years, and a large number of behavioural, neuropsychological and neuroimaging studies have significantly advanced our understanding of the developmental, cognitive and neural bases of face perception.

In this chapter we start our voyage towards the understanding of the features of human face recognition and we will learn about the most significant results on the cognitive and neural aspects of this fascinating topic. In particular, we start by analyzing this ability in normal subjects that do not show any sign of neurologic, psychiatric or neuropsychological disorder. Since it is important to know the normal features of a specific cognitive domain before we can start to comprehend the deficits, we will begin to address prosopagnosia, the disorder in face recognition, only in the next chapter.

Hopefully, after reading this chapter you would be able to answer important questions like: "Do faces represent a "special" category of stimuli for our visual system?", "How does face processing take place?", "Which is the neural underpinning of face processing?", "What is the speed of face processing?" and again, "Are face processing skills heritable?"

D. Rivolta, *Prosopagnosia*, Cognitive Systems Monographs 20,
DOI: 10.1007/978-3-642-40784-0_2, © Springer-Verlag Berlin Heidelberg 2014

2.1 Do Faces Represent a "Special" Category of Stimuli for Our Visual System?

2.1.1 Domain-Specific Hypothesis

In our environment we are surrounded by many different classes of visual stimuli such as cars, tables, chairs, churches, bottles, shoes, bodies and so on. The process of these objects (like all the other objects that you can have in mind) is mediated by *featural mechanisms*. This means that, put in simple words, we can recognize an object by combining all its features (tyres, windows, steering wheel) together (it is a car!).

Can we use the same featural mechanisms to process faces? In other words, can we recognize a face by putting all its features together? The answer is yes; we can do it. We can focus on the eyes or the nose of a person in isolation and recognize their "owner". However the evolution, over thousands of years, also equipped us with *holistic processing*.[1] Holistic processing enables us to perceive a face as a gestalt (a whole), which is more than the sum of the individual components. There is mounting empirical evidence showing the existence on holistic processing and researchers are starting to understand that holistic mechanisms are important for typical face processing. Below I will describe some of the experiments that (indirectly and directly) demonstrated holistic processing for faces. In particular, I will describe the face-specific effects that these experiments have shown.

Yin (1969) described for the first time what has been referred as the *face-inversion effect*.[2] This effect indicates that, in experimental environments, if people have to learn and remember faces they have never seen before (i.e., unfamiliar faces) they would be 20–25 % better at doing it when faces are shown upright than upside down (see Fig. 2.1).

Of course, you may think, that this is true even for objects. In fact we do not live in an upside-down world and inversion would affect the recognition of objects as well. This is true; the inversion effect occurs even for objects, but it is *much smaller* (up to 8 %) than the one shown for faces (see a very informative review on the topic in McKone et al. 2009). Yin suggested that this disproportionate inversion effect for faces is due to the fact that extracting the correct relationship between the face parts (holistic processing) was particularly important to face recognition and that extracting this information from inverted faces was difficult. In summary, according to this first key experiment, holistic processing does not "work" for inverted faces and it is not critically involved in object recognition.

[1] Over the past 20 years different authors have used different names such as configural, second-order relations or global to refer to what I define here as holistic processing. The theoretical reasons behind this go beyond the aim of the book.

[2] Also known as the *disproportionate inversion effect for faces.*

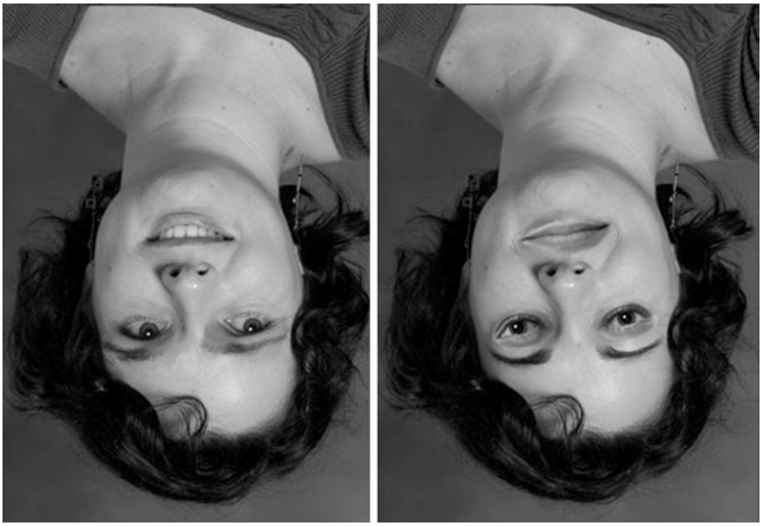

Fig. 2.1 It is difficult to recognize faces when they are shown *upside-down*. This picture represented here shows how a grotesque alteration of a face (try to turn the book *upside-down*) almost gets unnoticed on an inverted face (picture courtesy of Dr Rachel Robbins)

Tanaka and Farah (1993) provided more direct evidence in favor of holistic processing. They prepared a task (see Fig. 2.2) where people had to learn different identities such as "Tim" and were subsequently asked to recognize some of his features (e.g., the nose) in a forced choice task where Tim's nose had to be distinguished from another person's nose (e.g., Bob). Sometimes the two noses were shown in isolation and subjects had to indicate which one was Tim's nose (Fig. 2.2, top); sometimes the noses were shown within a face (Tim's nose in Tim's face and Bob's nose in Tim's face, see Fig. 2.2, bottom). The task was to indicate whether Tim's nose was on the left or on the right of the computer screen. Results indicated that the identification is better when features are shown within the face than when they are in isolation. This is called the *part-whole effect*. Importantly, this effect disappears (or is much smaller) in upside-down faces and with objects. Once again this effect suggests that upright faces only can benefit from holistic processing, since the face contour dramatically cues the correct recognition of a feature (nose) in it (McKone et al. 2009).

Another effect psychologists are very familiar with is the *composite-face effect*. This effect can be demonstrated in tasks where participants have to indicate whether, for example, the top-part of two sequentially presented unknown face stimuli is the same or different. Stimuli are shown in two different conditions: aligned or misaligned (see Fig. 2.3).

Results from different studies indicate that people are faster and more accurate when the halves are misaligned. This is because aligned faces (even when made up of two different identities) automatically "glue-up" to form a new configuration

Fig. 2.2 The part-whole effect demonstrates that the identification of a particular feature (e.g., eyes) is facilitated when this is presented within the face configuration (*top*) then when in isolation (*bottom*). Figure obtained with permission from Palermo and Rhodes (2002)

Fig. 2.3 The composite-face effect. Judging whether the *top halves* of two pictures depict the same person (as in this figure) is harder when the two halves are aligned (forming a new identity) than when they are misaligned (these face stimuli were kindly provided by Prof. Daphne Maurer from McMaster University, Canada)

(a new identity). It is then extremely difficult to process one half of the face without being influenced by the perception of the other half of the face, which makes the task more difficult. Once more, this effect is absent for other objects and for upside-down faces, strengthening the dedicated role played by holistic mechanisms in upright face perception (Robbins and McKone 2007; Young et al. 1987).

The last approach I wish to describe for the assessment of holistic processing uses artificially modify faces to change the spacing between features (for example the distance between the eyes) or the features themselves. In a very well-known

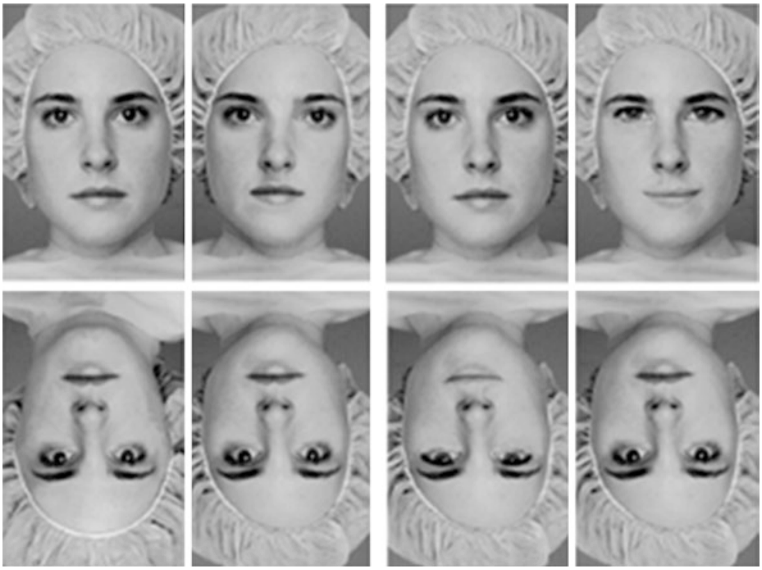

Fig. 2.4 Examples of stimuli used in the Jane task. It is harder to detect the differences in the "spacing condition" (*top* and *bottom left*) than in the "features" condition (*top* and *bottom right*) when faces are upside down (these face stimuli were kindly provided by Prof. Daphne Maurer from McMaster University, Ontario, Canada)

task called the "Jane task" (Le Grand et al. 2004; Mondloch et al. 2002) participants had to determine whether two sequentially shown faces where the same or different, when the difference was either in the spacing of the features or in the features themselves (see Fig. 2.4).[3]

The same task is then given with upside-down stimuli. Results showed that performance decreases, overall, for upside-down stimuli, but that this inversion affects the detection of changes to spacing more than the detection of changes to features. Since, as we should know by now, face inversion is supposed to affect holistic processing, this result on the Jane task is in agreement with the idea that holistic mechanisms process the spacing between face parts and not the parts themselves.[4]

What we have learned so far is that faces, inverted faces and objects are processed by *featural mechanisms*. However, the perception of upright faces is mediated by *holistic mechanisms* more than any other type of stimuli. The effects reported above strongly support this claim and, as such, support what is known

[3] In the original Jane task there is also a third condition, called the "contour condition", that for clarity reasons I do not report here. The interested reader is invited to refer to the original articles.

[4] Some researchers believe that holistic processing involves even the features of faces, but it is beyond the aim of the book to address those theoretical issues. See McKone and Yovel (2009) for a detailed description of the issue.

as the *domain-specificity hypothesis*. The domain-specificity hypothesis states that faces represent a “special” category of stimuli that are processed by holistic mechanisms (McKone et al. 2006). Without these mechanisms our recognition would be problematic, as we will see in the next chapter. You may think that the experiments I showed you above lack of *ecological validity* that is, our visual system is never exposed to those stimuli in everyday life. This is in part true; how often would you see a misaligned composite face at the pub? However, in order to study our cognition we need somehow to decompose our complex processes (e.g., visual cognition) into smaller and investigable units. This is the reason why each of the tasks I presented above focuses only on one specific aspect of face processing. In addition, this process of division in subtasks enables us to have a rigid control over the variables under investigation.

2.1.2 The Expertise Hypothesis

There are researchers that do not support the domain-specificity hypothesis. They support the *expertise hypothesis*; which states that expertise plays a critical role in developing holistic mechanisms for faces. In general, the basic difference between face and object perception is the “depth” of processing: when we see a face we can identify it, whereas when we see a table we usually do not. According to this view, experts in a particular field such as dog experts (e.g., people who can individually identify many different individual Golden Retrievers) or car experts (i.e., people that can identify, in a glance, many different makes of cars) should show holistic processing not only for faces but even for their category of expertise. Even though some early studies supported the expertise hypothesis, more recent and better controlled experiments failed to support the expertise hypothesis, suggesting that only upright faces rely on holistic processing (McKone et al. 2006, 2009).

Let me give you some examples of the evidence discarding the expertise hypothesis (experiments reported in Robbins and McKone 2007). In these experiments the authors tested dog-experts (people with, on average around 23 years of experience as dog judges, breeders or trainers) and novices, that is, non-experts in dog recognition. In theory, both experimental groups are expert in face perception, but only dog-experts are also expert in dog perception. Accordingly, following the expertise hypothesis, this manipulation should demonstrate holistic effects not only in face processing but also in dog perception for dog-experts only. In one of these experiments dog experts and novices had to memorize (upright) dogs and faces. After an interval of few minutes where they had to complete a different task, the memorized stimuli were shown on the screen with a distractor (i.e., a stimulus that was not shown during the learning phase); this means that, for each trial, each subject hat to decide whether the learned stimulus was on the left or on the right of the computer screen. The procedure was repeated when stimuli were shown upside-down (rotated 180°). Results demonstrated that, as expected, novices showed a face inversion effect; that is, the inversion affected more the memory

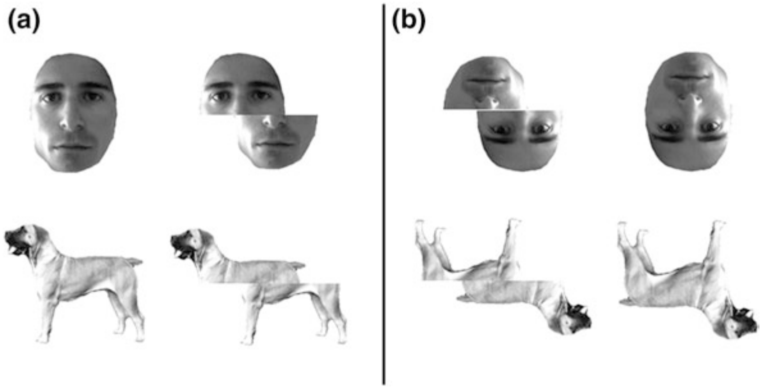

Fig. 2.5 Facsimile of stimuli adopted in the composite-face task. Faces and dogs were shown in the aligned and misaligned condition both (**a**) upright and (**b**) upside-down (These stimuli were not the actual one adopted in the original experiments. I created them for representational purposes only)

performance for faces than for dogs. Interestingly, despite their great experience with dogs, even dog-experts showed bigger face inversion than dog inversion effects. This strongly supports the domain-specificity hypothesis (i.e., faces are special and processed by holistic processing) and not the expertise hypothesis (i.e., all stimuli with which subjects had strong experience can be processed like faces, using holistic processing).

Another experiment addressed whether the composite face effect could be demonstrated with stimuli of expertise (i.e., dogs) in dogs-experts. The procedure is similar to the composite-face task I presented above. As in classical composite-face tasks, the faces were shown aligned and misaligned, in both upright and inverted conditions. The same happened for dogs, they were presented aligned and misaligned stimuli, both upright and upside down. Results demonstrated that novices showed, as expected, the classical result: they were more accurate when the halves were misaligned then when they were aligned, that is, they showed a composite-face effect (see Fig. 2.5 for a description of the results). Similarly to many other studies, results showed that this effect disappeared when stimuli were shown upside-down. Note that in both orientations they did not show any composite effect for dogs; this was expected since novices had no experience in dog recognition whatsoever. Will dog experts show the composite effect for dogs, the class of stimuli they have a lot of expertise with? Results were negative; dog-experts showed a composite effect for faces only and not for dogs. Similar to novices, they did not show any composite effect for inverted stimuli (Fig. 2.6).

Overall these two experiments, along with many others, demonstrate that holistic mechanisms are features of upright faces processing only. This is not because we acquire expertise with faces during our lives, but because, as we will see in the

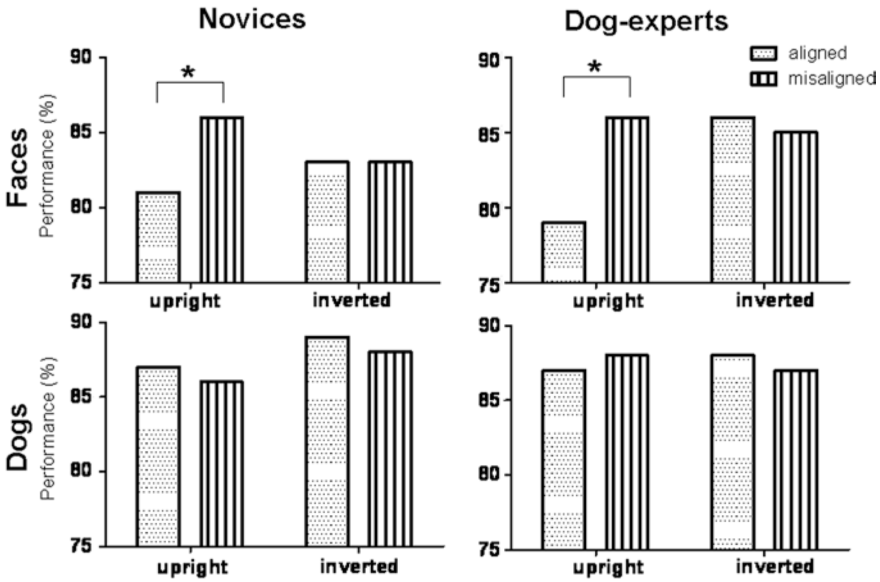

Fig. 2.6 Results on the composite task. Both novices (*left column*) and dog-experts (*right column*) show a composite effect for faces (the *asterisk* indicates a statistically significant difference between the two conditions). No group shows a composite effect for dogs, suggesting that expertise with a category of stimuli does not develop holistic mechanisms for that category (This figure is not taken from the original work of Robbins & McKone (2007), but it only wishes to represent, with fictitious data, the main effects described in the original work)

next paragraphs, we might be born with this ability. Thus, the answer to the question: "Do faces rely on "special" cognitive processes?" is *yes*. The processing of upright faces only relies on dedicated cognitive mechanisms. This makes faces a *special* stimuli for the human visual system.[5]

2.2 How Does Face Recognition Take Place?

It takes us a small fraction of a second to recognize a face once we perceive it. Even though we experience every day the speed of our face processing system, it is known that the recognition of a familiar face does not occur all at once, but involves different cognitive steps. In 1986 two British researchers (Vicki Bruce and Andy Young) described those steps in an influential cognitive model of face processing known as the "Bruce and Young model of face recognition" (Fig. 2.7). The model was developed to accommodate a range of empirical observations from

[5] The reader interested in the theoretical debate between the competing domain-specific and expertize hypothesis is invited to read works from Gauthier and colleagues.

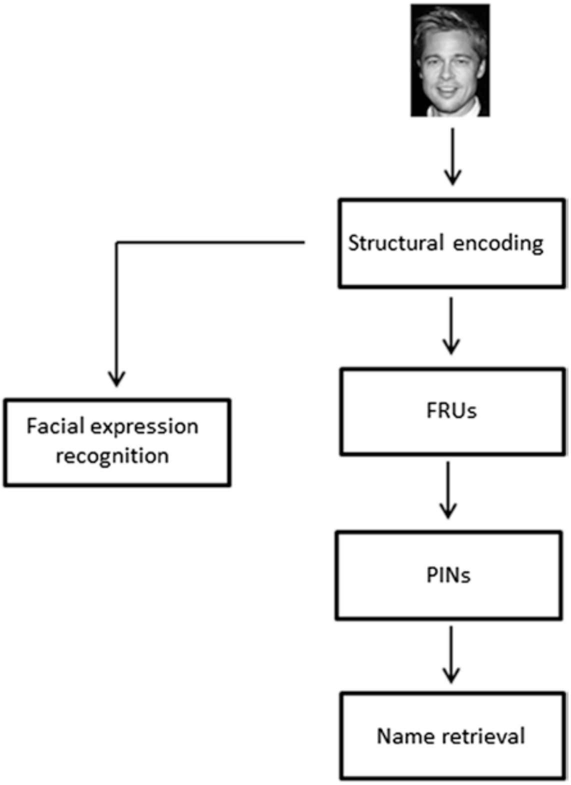

Fig. 2.7 A simplified version of the Bruce and Young (1986) model of face recognition. This model depicts the stages between the perception of a face and its recognition (*FRU*s face recognition units; *PIN*s person identity nodes; see text for a complete description of the model)

normal subjects and neuropsychological patients, and proposed an organization of the cognitive face processing system that is hierarchical and branching.

The model proposed that face recognition occurs via an initial stage of *Structural Encoding* of a face's appearance, where both holistic and featural processing are supposed to occur (Schmalzl 2007). This is followed by firing of *Face Recognition Units* (FRUs) that respond to particular features and configurations of specific familiar faces, access to relevant semantic information within the *Person Identity Nodes* (PINs), and finally name retrieval. In addition, it proposed that this process of person identification occurs independently from the identification of facial expressions and lip reading. Bruce and Young postulated that the representations of identity and those of the more changeable aspects of a face (i.e., emotion expression) must be (at least to some extent) independent from one another, or else a change in expression or a speech-related movement of the mouth could be misinterpreted as a change of identity.

With an example, we can see that Brad Pitt's face is processed within the structural encoder that codes for the structure of the face such as having two eyes above the nose above the mouth, and the distance between them (we will describe in more detail this in the next paragraph). This information is fed to the FRUs that "fire" because the face is familiar and then to the PINs where biographical information such as that 'this person was married to Jennifer Aniston, is currently (January 2013) married to Angelina Jolie, starred in Troy, Fight Club, The curious case of Benjamin Button' and so on. Finally we can retrieve the name: "Brad Pitt!" According to the model, the understanding of whether the face of Brad Pitt depicted in the photo looks happy, angry, sad or disgusted is mediated by a different system. In Chap. 3 we will see how this model can account for problems in face perception.

2.3 What is the Neural Underpinning of Face Processing?

When you look at an object from the outside world, the light reflects from it and bounces into our eyes, where a structure called the retina decodes this information and passes it to the brain. As discussed in Chap. 1, the occipital lobe at the back of our brain is the main anatomical region that receives this information. From the occipital lobe, information is diverted to other brain regions such as the temporal lobe and the parietal lobe (Milner and Goodale 2006) (Fig. 2.8).

We have known for a few decades that the temporal lobe is crucial for our ability to recognize faces, objects and places. Let's focus now on faces. The first question here is: "How can we know about the crucial role played by the temporal lobe in face recognition?" Single unit recordings (the technique that consists of directly placing electrodes in the brain, through the skull, and measuring the firing rates[6] of neurons) in anesthetized monkeys paved the way for the current knowledge we have in face recognition. In 1969 Gross and colleagues recorded the activity of neurons in the inferior part of the macaque temporal cortex. They found that there were groups (patches) of neurons that responded virtually *only* when the animal was looking at faces and not when it was looking at fruits, hands or tools. Some of these neurons responded to front view faces, other to profile faces, other to face parts (i.e., eyes). There were even single neurons that responded to some specific identities (people that monkeys were familiar with such as the researchers working with them). These results collected over 50 years ago strongly suggested that the face recognition system, at least in monkeys, has developed some modules that process faces only, and they provided the first anatomical substrate that underlies the special status that faces play for social functioning (see a review on the topic in Gross 2008).

Only after many years, with the advent of modern functional neuroimaging techniques such as positron emission tomography[7] (Haxby et al. 1994) and fMRI, researchers have attempted to localize face recognition processing within the

[6] Firing rates are typically measured as the number of action potentials (*spikes*) a neuron fires in 1 s.

[7] PET belongs to the class of invasive neuroimaging techniques. This technique enables to see brain activity only after the intravenous injection of a radioactive substance. PET represented one of the most adopted techniques for the visualization of brain activity in vivo before fMRI was invented.

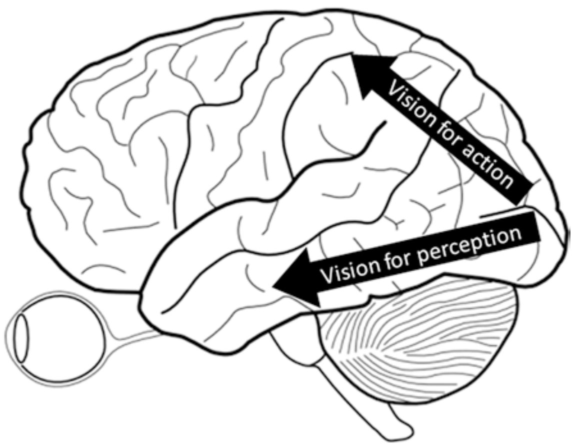

Fig. 2.8 Lateral view of the left hemisphere. The *two arrows* indicate the two visual systems: the one we use for perception (*vision for perception*) and the one we for action (*vision for action*). The first, also known as occipito-temporal route, mediates objects and face recognition; the second one, also known as occipito-parietal route, mediates visually guided movements, such as reaching a glass

human brain. One of the most influential findings was the discovery made by Kanwisher and colleagues (1997). By using fMRI the authors identified and localized the so called Fusiform Face Area (FFA) within the human temporal lobe. In this well-known experiment people were shown faces, cars, hands and other objects, while their neural (BOLD) activity was recorded. Results showed that, similarly to findings in monkeys, there was a region (the FFA) within the human temporal lobe that showed a response for faces that was at least twice as strong as for other objects. This seminal research was subsequently confirmed in many other studies leading to other research questions such as whether the FFA represents the region where holistic processing, the face-dedicated processing, takes place. It is now believed that FFA represents a crucial region for holistic processing, since the FFA responds stronger for upright face processing than inverted face processing[8] and it mediates the composite face effect (remember that face inversion and the composite face effect are believed to demonstrate holistic processing) (Liu et al. 2009; Schiltz and Rossion 2006; Yovel and Kanwisher 2005). In addition to holistic processing, however, it has been shown that the activity of FFA is also involved in face identification and recognition (Rotshtein et al. 2005). It should be noted however that further research failed to confirm the role of FFA in face identification (Kriegeskorte et al. 2007) and further studies are needed to clarify the issue.

Over the last 20 years, numerous cortical face-sensitive brain regions have been discovered in humans. Each of them seems to represent the neural correlate of different behavioural phenomena. The occipital face area (OFA, Gauthier et al. 2000) in the occipital lobe seems to respond mainly to face features, the superior temporal sulcus (STS) to changeable aspects of the face such as facial expression, and the anterior temporal face patch (ATFP) of the anterior temporal lobe to face identities (Gobbini and Haxby 2007; Haxby et al. 2000; Kriegeskorte et al. 2007) (Fig. 2.9).

[8] Face inversion increased the activity of object selective regions, further suggesting that inverted faces are processed using mechanisms that are common to object processing.

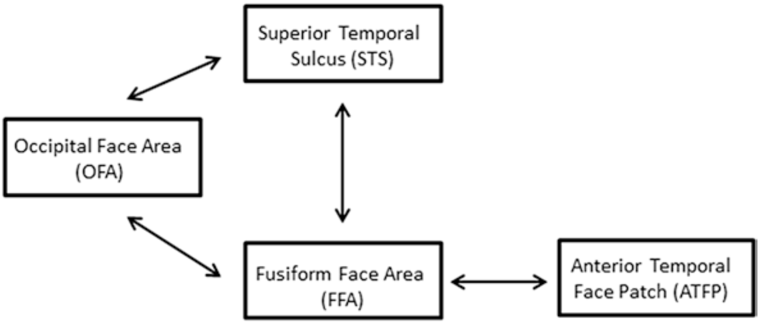

Fig. 2.9 Simplification of the Haxby et al. (2000) neural model of face processing. This model schematically represents the main human neural substrates for normal face processing. The OFA may be the first "gate" of the system, where a stimulus is judged to be a face and where its features are processed. This information bifurcates to the STS for the process of emotions (and other changeable aspects of faces such as eye gaze) and to the FFA, where holistic processing (and maybe identification) is carried on. The identification of faces occurs in (but not only) the ATFP

Causal evidence for the crucial role played by OFA in the processing of face features comes from the use of TMS, which, as seen in the previous chapter, temporarily "deactivates" a specific brain region and monitors the consequent behaviour. Pitcher and colleagues (2007) presented participants with faces and houses that changed in the shape of their features (e.g., different eyes, different windows) or in the spacing of their features (i.e., the distance between features changes, but the features themselves remained the same). As the reader may remember this manipulation is similar to the one described for the Jane task above. The idea behind the manipulation of this study is that the spacing-change detection reflects holistic process, whereas feature-change detection reflects features processing. Participants had to determine whether two sequentially shown pictures of faces or houses depicted the same stimulus or not (that is, if there was a spacing change or a feature change or no change). Results demonstrated that TMS delivered within 60–100 ms post-stimulus onset on OFA (previously determined for each subject individually using fMRI) disrupted the face-performance on the feature part of the task, and not the spacing task. The specificity of this disruption for face processing was demonstrated by showing the total absence of TMS stimulation consequences on house discrimination. In summary, this study showed the critical role played by OFA in the processing of face features; so OFA does not seem to represent the locus of holistic face processing.

The importance of STS for changeable aspects of a face was given by Narumoto and colleagues (2001). The authors, amongst other conditions, asked participants to place their attention to the expression or the identity of faces they were shown. Results obtained with fMRI indicated that the (right) STS was more involved in the coding of facial expression than identity, suggesting the involvement of STS in changeable aspects of the face (and not in identity discrimination). The role of the ATFP for face identification has recently been given, using fMRI, by Kriegeskorte and colleagues (2007). Authors demonstrated that the pattern of fMRI activity within this region, but not FFA, could discriminate between two different identities participants had to learn. Overall, the differential role played by these face areas in humans has until now not definitely been clarified and future research will address the issue (in Chap. 3 we will see how these face regions are involved in disorders of face recognition). What seems clear from previous research in cognitive science is that the right hemisphere of the human brain is dominant for face processing. Dominant means that great part of the work in face recognition is done by the right hemisphere.

The existence of face specific clusters of neurons in the human brain has not only been determined by fMRI, but it has been very often confirmed even by (invasive) single unit recordings in patients with drug-resistant epilepsy. These patients cannot find relief from their condition with the use of anti-epileptic drugs and they sometimes need surgery aimed at the removal of the neural tissue generating the epileptic seizures (*focus epilepticus*). In order to remove the correct region, some invasive recordings are necessary and researchers have the unique opportunity to run exceptional experiments aimed at clarifying the relation between the mind and the brain (Duchowny 1989).

In 2005 it was discovered the existence of neurons in the human (medial) temporal lobe that respond preferentially to some specific identities. For example authors described a neuron that fired with Jennifer Aniston's (a famous American actress) pictures only. Different pictures of the actress taken many years apart with different hairstyle and different orientation of the face elicited the response of these neurons. Importantly those neurons did not respond to faces of other celebrities. Other groups of neurons responded only to other famous people such as Michael Jordan (a famous Basketball player) or Bill Clinton (a former US president). Furthermore, recent evidence in two cases of people with drug-resistant epilepsy demonstrated that the electrical stimulation of face-sensitive regions within the fusiform gyrus (Parvizi et al. 2012) and within the lateral occipital gyrus (Jonas et al. 2012) completely distorted face recognition, thus leading to "transient prosopagnosia".

In summary, there is a convergence from animal and human studies that the face recognition system relies on dedicated neural populations. This, once again, suggests that faces are "special"; even at the neural level. What happens at the behavioural level when the brain regions involved in face processing get damaged will be the topic of the next chapter.

2.4 What is the Speed of Face Processing?

2.4.1 Neurophysiological Investigations

As we have just learned, there exist regions within the human brain that are specifically "tuned" for face processing. However, we still do not know how long it takes for these regions to compute the processes. Are these regions fast at processing? How fast?

As explained in Chap. 1, EEG and MEG (but also TMS) are the best available non-invasive techniques to adopt for answering this question. Since the late 80's it is well known that the existence of an electrophysiological response for faces that has been recorded both with intracranial invasive methods in people with drug-resistant epilepsy (Allison et al. 1994; Allison et al. 1999) and with surface EEG (Bentin et al. 1996; Jeffreys 1989). When people are shown faces, their EEG activity shows a negative deflection occurring at around 170 ms (ms) post stimulus onset; this is known as the N170. The N170 is detectable from occipito-temporal surface electrodes and consistently shows bigger amplitude for faces than other categories of visual stimuli. It is believed that the N170 is mainly generated by the activity of two cortical regions such as the occipital lobe (OFA) and the temporal lobe (FFA), where there are neurons "tuned" for faces (Deffke et al. 2007; Itier et al. 2007; Linkenkaer-Hansen et al. 1998). Since these brain regions have face-sensitive neurons, it is believed that the synchronous firing of thousands of those neurons can give rise to a potential that is big enough to be seen on the scalp (N170). As expected, the face-sensitive activity at around 170 ms post stimulus onset has been detected even with MEG and it is named M170 (Liu et al. 2000). Albeit, as we just said, the generators of the N/M170 are two, it is still largely unexplored whether each of these two regions can generate a distinct N/M17. In other words: "Can the two sources of the N/M170 code for a different kind of face processing?" I will answer this question in the next chapter when I describe congenital prosopagnosia. For now we focus on a classical finding. The peak of the N/M170 is delayed of 10–13 ms when faces are presented upside-down. This inversion effect of the N/M170 suggests that holistic processing occurs at around 170 ms post-stimulus onset (Bentin, et al. 1996; Rossion et al. 2000).

The N/M170 represents by far the most explored face-sensitive electrophysiological component. However we know that there is a face-sensitive component that peaks earlier than the M170; this component peaks at around 100 ms post stimulus onset and since it has been firstly investigated with MEG, it is called the M100 (Liu et al. 2002). This component is generated from the occipital lobe and supposedly codes for aspects of face processing distinct to the M170. One hypothesis is that M100 reflects the detection that a face is present in the visual field, whereas the N/M170 enables the identification of it (Liu et al. 2002). Other lines of evidence suggest that the M100 codes for face features, whereas the N/M170 is sensitive to holistic processing (Pitcher et al. 2011), albeit some evidence showed that the N/M170 codes for both features and holistic processing (Harris and Nakayama 2008). One of the main reasons for these discrepancies in results may rise from the different technique (EEG versus MEG), experimental design, methodology and data processing. One recent study shed further controversy on the topic by indicating that even the M100 can be sensitive to face familiarity (Rivolta et al. 2012). Future studies will hopefully clarify the issue.

2.4.2 Behavioural Investigations

This impressive speed our face recognition system shows for face processing further supports neuroimaging and behavioural data claiming that faces have a special status in our face recognition system. Since we already described the most important neuroimaging findings in face processing research, let's focus our attention on some behavioural experiments that attempted to understand how long it takes to determine that there is a face in a visual scene and how long it takes to identify a familiar face. By using a Rapid Visual Serial Presentation (RVSP) paradigm, characterized by the rapid and sequential presentation of visual stimuli, it has been shown that object categorization (i.e., deciding whether a visual stimulus is a face, an animal, or an object) occurs just as rapidly as the mere detection of an object within a visual field. Since the effect occurred even when stimuli were shown for only 17 ms on the screen, this result strongly suggests that object detection occurs as quickly as its categorization, thus indicating that stimulus detection and categorization may occur in parallel (Grill-Spector and Kanwisher 2005; Purcell and Stewart 1988). This result is in line with the common experience that as soon as we see something we can say that it is a face (or an object). Interestingly, in agreement with MEG results described above, it has even been demonstrated that face identification can occur in around 100 ms (Tanaka 2001). Further evidence supporting the exceptional speed of face processing comes from investigations that monitor the eye movements. Using a specific device, called the *eye-tracker*, it is possible to monitor the speed and the features of eye movements. In other words, it is possible to see where and for how long people focus their sight on visual stimuli presented on a computer screen. It has been recently shown that when people have to make eye movements towards target stimuli such as faces, animals and vehicles, they are on average more accurate and much faster when they have to do it for faces than other categories of visual stimuli. In addition, the minimum saccadic reaction time towards faces occurs in 110 ms, faster than for animals (120 ms) and vehicles (140 ms) (Crouzet et al. 2010).

2.5 Are We Born with "Face-Specific" Cognitive and Neural Mechanisms?

As described at the beginning of this chapter, over the past decades there has been ongoing debate about whether face specific cognitive mechanisms, that is holistic processing, is acquired or it is present since birth. One way to find the solution to this issue is by investigating face processing in very young kids; infants in particular. For practical reasons the methodologies adopted for research in infants are very different from the ones adopted for adults (would you expect a 4 day-old infant to perform the composite-face task?). Typically, *looking time* (i.e., the infant spends more time fixating something of interest than something less interesting) measures are adopted as an index of preference.

Research accumulated over the last 10 years strongly suggests that humans are equipped with face-specific cognitive mechanisms *from birth*. Until recently however, it was believed that children need around 10 years of experience with faces to show the face-specific experimental effects described above in adults (Carey et al. 1980). This, which was initially taken as strong support for the expertise hypothesis (Diamond and Carey 1986), has subsequently been demonstrated to be wrong (see McKone et al. 2012; McKone et al. 2009 for an excellent and in detail description of the issue).

2.5.1 Behavioural Studies in Infants

Behavioural studies have indicated that newborns show a preference for tracking face-like configurations compared to other types of visual arrays (Fig. 2.10). Three-month-olds (and even 3-days-old!) can recognize the identity of novel individuals with similar looking faces (presented without hair and across view changes), suggesting that a face representation tuned to upright faces and able to support individual level representations is present at birth (Pascalis et al. 1998; Slater et al. 2000; Turati et al. 2008). Newborns less than a week old prefer attractive over unattractive (as rated by adults) faces when stimuli are upright, but not when inverted (Slater et al. 2000).

In addition, children as young as 4 years of age can show an inversion effect, a part whole effect, a composite effect, as well as general sensitivity to spacing of face features (for example as tested on the Jane task) (Cohen and Cashon 2001; Hayden et al. 2007; McKone et al. 2009). In addition, these effects not only are presents in young kids, but (by around 5 years of age) also show similar magnitude of the effects seen in adults. That is, the effect is not present, but has the same size as the one observed in adults (McKone et al. 2012). In summary, so far, all behavioural findings of adults' face processing are present with similar strength in infants. Thus, developmental studies strongly suggest that years of experience and extensive practice with face individuation cannot be the only factor that accounts for the observation of face specific experimental effects (i.e., the expertise hypothesis cannot be supported).

Furthermore, all infants (and even monkeys) show *perceptual narrowing*. It has been demonstrated that infants have, at birth, the capacity to represent all faces (faces from all races and even monkey faces!). This ability narrows down during time and infants become specialized for types of face that they are exposed to frequently in their native environment. Results showed that 6-month-old infants could discriminate both humans and monkey faces, while 9-months-old and adults could only discriminate human faces (Pascalis et al. 2002). Perceptual narrowing for faces explains why we tend to be better at perceiving, memorizing and identifying people from our own race than other races. This effect is called the

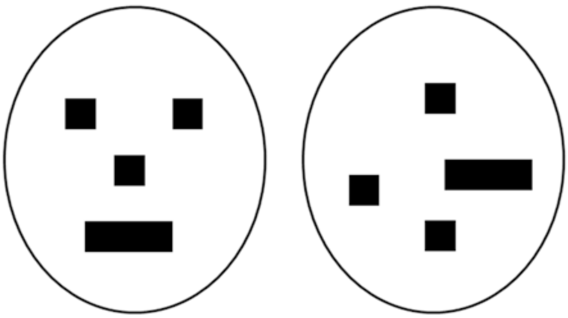

Fig. 2.10 Research with newborns has adopted stimuli like these. On the *left* there is a face-like configuration, whereas on the *right* there is a configuration that, despite showing the same elements, does not resemble a face. Newborns spend more time looking at the stimulus on the *left* than the one on the *right*

other-race effect (Feingold 1914). If you are European and grew up in Europe, when travelling to Asia for the first time you may find it more difficult to differentiate between Asian people than you do for people in your own country. The same is true, of course, for Asians that come to Europe for the first time. Perceptual narrowing for faces is similar to what happens in the domain of language (Kuhl et al. 2003), where newborns can discriminate phoneme boundaries from all possible languages (i.e., we can potentially learn all languages at birth), but they lose this ability with time. At first the perceptual narrowing for faces seems a negative aspect of development, since we are losing something that might be useful; however, this specification toward a particular race enhances our ability to discriminate people within this race.

The last aspect I wish to underline is the idea of the existence of a *critical period* (Sengpiel 2007). Experimental evidence suggests that face processing is characterized by a critical period, requiring adequate environmental input (i.e., normal face perception) before the face-system is used for other purposes. What happens if we do not show faces to infants for a few months (or years)? Of course we are not so cruel as to do this to our kids, but children with *congenital cataracts* in the eye/s do not receive proper visual inputs until the cataracts get removed (usually around 1 or 2 years old). Le Grand and colleagues (2004) demonstrated that even many years after the removal of the cataracts (even 12 years as indicated in Ostrovsky et al. 2006), years in which they had normal exposure to faces, people showed impaired holistic processing (e.g., impaired performance on the composite-face task and on the Jane task).

In summary, behavioural data reviewed above demonstrating adult-like face abilities present at birth, perceptual narrowing and critical periods, are all consistent with a genetically determined *innate* contribution to infant face recognition.

2.5.2 Neuroimaging Studies in Infants

A second line of evidence supporting the innate origin of our face recognition abilities comes from neuroimaging investigations. Many studies have shown, as described above, that a region in the temporal cortex, the FFA, responds more robustly to faces than to other types of visual stimuli. Similarly, studies using M/EEG indicated that brain responses occurring approximately 170 ms after stimulus onset typically show greater amplitude for faces than for other types of objects. When tested using fMRI, around 80–85 % of children between 5 and 10 years of age showed an FFA. In developmental studies, the size of (right) FFA correlated with behavioural performance on face memory task but not with object memory. In addition, ERPs studies in young children reported a face-specific N170 that showed even a face-inversion effect (Scherf et al. 2007).

Altogether the results reviewed so far demonstrate that both behavioural and neuroimaging findings reported in adults can be found in the youngest age group tested. However, even if the size of behavioural effects and the N170 seem to be comparable with adults, the magnitude of FFA activity is less clear and will require further research. Overall, research seems to support the claim that humans are born with innate face-specific cognitive and neural mechanisms.

2.6 Are Face Processing Skills Heritable?

It is well known in psychology that people show variability in their performance on different cognitive tasks. This variation represents the base of *individual differences*. Some of us are very good at calculations other are less good; some people have very good memory skills, others do not, and this also goes for reading abilities, athletic performance, musical skills, etc. Are we variable in face recognition abilities or everyone has the same skills? Similar to other abilities, many studies have underlined the existence of strong individual differences in face processing (Bowles et al. 2009). A key question regarding the understanding of whether this variability in face processing skills is "genetic", heritable, or whether the environment (i.e., the exposure to faces, social status of the family, country of origin and so on) plays a critical role in shaping our skills.

It is known that general intelligence is highly heritable in humans. Can, however, a specific ability such as face recognition show an high level of heritability? In other words, can someone inherit high intelligence, but poor face recognition skills from the parents? (McKone and Palermo 2010). Twin studies constitute an interesting methodological approach to the issue. Twins can be monozygotic or dizygotic. Monozygotic twins share the 100 % of their genes, whereas dizygotic twins share around 50 %. This means that, since twins usually share the same familiar environment, a difference in the correlation between specific measures (i.e., face memory) in the two groups must be attributed to genetic and not environmental factors.

Wilmer and colleagues (2010) looked at face memory skills in 164 monozygotic and 125 same gender dizygotic twins. Results showed that there is a correlation between performances of monozygotic, but not for dizygotic twins; this means that if twin A of a monozygotic twin couple performed very well, twin B tended to do the same. On the other side, if twin A of a dizygotic twin couple performed very well, performance of twin B was not necessarily also good. Since monozygotic and dizygotic twins share the same environment, this difference posits for genetic factors in individual variations in face processing. The question of which genes are involved however remains unanswered. Importantly, both monozygotic and dizygotic twins were not correlated in their performance on a memory task that did not tap into face processing such as abstract art memory or a paired-associates memory test, indicating that face processing skills *only* are heritable and they do not depend on general attention and/or memory functioning. Overall these data add to previous results strongly indicating special cognitive and neural mechanisms for face processing. In the next chapter we will further discuss the role of familiarity in face processing skills by discussing cases of congenital prosopagnosia.

2.7 Conclusions

In this chapter we have learned that upright face processing only is mediated by specific cognitive (i.e., holistic mechanisms) and neural (e.g., OFA and FFA) mechanisms. In addition, face processing is mandatory, occurs very quickly and it is mediated by face-sensitive physiological mechanisms (e.g., M100, M170). All these features seem to be present from birth and not acquired (although they may be improved) with experience. Overall, the evidence reviewed in this chapter strongly indicates that faces represent *special* stimuli for our visual system. In other words, faces seem to represent the category of visual stimuli that engage the fastest and most dedicated cognitive and neural processing.

In the next chapter we will learn that this special and precious ability most of us share can fail, causing serious and embarrassing problems in face recognition. Thus we will talk about people that have lost their ability to recognize faces after brain injuries (*acquired prosopagnosics*) and about people that have never developed the typical ability to recognize face (*congenital prosopagnosics*). We can anticipate that since prosopagnosics typically have specific problems in face processing while their object processing is spared (or much less impaired than face processing), the existence of prosopagnosia further supports the special role played by faces in humans.

References

Allison, T., McCarthy, G., Nobre, A., Puce, A., & Belger, A. (1994). Human exstrastriate visual cortex and the perception of faces, words, numbers, and colors. *Cerebral Cortex, 5*, 544–554.

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*, 415–430.

Bentin, S., McCarthy, G., Perez, E., Puce, A., & Allison, T. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience, 8*, 551–565.

Bowles, D. C., McKone, E., Dawel, A., Duchaine, B., Palermo, R., Schmalzl, L., et al. (2009). Diagnosing prosopagnosia: Effects of aging, sex, and participant-stimulus ethnic match on the Cambridge Face Memory Test and Cambridge Face Perception Test. *Cognitive Neuropsychology, 26*(5), 423–455.

Bruce, V., & Young, A., (1986). Understanding face recognition. *British Journal of Psychology, 77* (3), 305–327.

Carey, S., Diamond, R., & Woods, B. (1980). Development of face recognition—a maturational component? *Developmental Psychology, 16*(4), 257–269.

Cohen, L. B., & Cashon, C. H. (2001). Do 7-month-old infants process independent features of facial configurations? *Infant and Child Development, 10*, 83–92.

Crouzet, S., Kirchner, H., & Thorpe, S. J. (2010). Fast saccades toward faces: face detection in just 100 ms. *Journal of Vision, 10*(4), 1–17.

Deffke, I., Sander, T., Heidenreich, J., Sommer, W., Curio, G., Trahms, L., et al. (2007). MEG/EEG sources of the 170-ms response to faces are co-localized in the fusiform gyrus. *NeuroImage, 35*, 1495–1501.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General, 115*(2), 107–117.

Duchowny, M. S. (1989). Surgery for intractable epilepsy: Issues and outcome. *Pediatrics, 84*, 886–894.

Feingold, C. A. (1914). The influence of environment on identification of persons and things. *Journal of Criminal Law and Police, 5*, 39–51.

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.

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

Grill-Spector, K., & Kanwisher, N. (2005). As soon as you know it is there, you know what it is. *Psychological Science, 16*(2), 152–160.

Gross, C. G. (2008). Single neuron studies of inferior temporal cortex. *Neuropsychologia, 46*, 841–852.

Harris, A., & Nakayama, K. (2008). Rapid adaptation of the M170 response: Importance of face parts. *Cerebral Cortex, 18*, 467–476.

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

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.

Hayden, A., Bhatt, R. S., Reed, A., Corbly, C. R., & Joseph, J. E. (2007). The development of expert face processing: Are infants sensitive to normal differences in second-order relational information? *Journal of Experimental Child Psychology, 97*, 85–98.

Itier, R. J., Alain, C., Sedore, K., & McIntosh, A. (2007). Early face processing specificity: It's in the eyes! *Journal of Cognitive Neuroscience, 19*(11), 1815–1826.

Jeffreys, D. A. (1989). A face-responsive potential recorded from the human scalp. *Experimental Brain Research, 78*, 193–202.

Jonas, J., Descoins, M., Koessler, L., Colnat-Coulbois, S., Sauvee, M., Guye, M., et al. (2012). Focal electrical intracerebral stimulation of a face-sensitive area causes transient prosopagnosia. *Neuroscience, 222*(11), 1078–1091.

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

Kriegeskorte, N., Formisano, E., Sorger, B., & Goebel, R. (2007). Individual faces elicit distinct response patterns in human anterior temporal cortex. *Proceedings of the National Academy of Science USA, 104*(51), 20600–20605.

Kuhl, P. K., Tsao, F.-M., & Liu, H.-M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences (PNAS) USA, 100*, 9096–9101.

Le Grand, R., Mondloch, C., Maurer, D., & Brent, H. (2004). Impairment in holistic face processing following early visual deprivation. *Psychological Science, 15*, 762–768.

Linkenkaer-Hansen, K., Palva, J. M., Sams, M., Hietanen, J. K., Aronen, H. J., & Ilmoniemi, R. J. (1998). Face-selective processing in human extrastriate cortex around 120 ms after stimulus onset revealed by magneto- and electroencephalography. *Neuroscience Letters, 253*, 147–150.

Liu, J., Harris, A., & Kanwisher, N. (2002). Stages of processing in face perception: An MEG study. *Nature Neuroscience, 5*(9), 910–916.

Liu, J., Harris, A., & Kanwisher, N. (2009). Perception of face parts and face configurations: An fMRI study. *Journal of Cognitive Neuroscience, 22*(1), 203–211.

Liu, J., Higuchi, M., Marantz, A., & Kanwisher, N. (2000). The selectivity of the occipitotemporal M170 for faces. *Cognitive Neuroscience and Neuropsychology, 11*(2), 337–341.

McKone, E., Crookes, K., Jeffery, L., & Dilks, D. (2012). A critical review of the development of face recognition: Experience is less important than previously believed. *Cognitive Neuropsychology, 29*, 174–212.

McKone, E., Crookes, K., & Kanwisher, N. (2009). The cognitive and neural development of face recognition in humans. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences IV* (pp. 467–482). Cambridge, MA: Bradford Books.

McKone, E., & Yovel, G. (2009). Why does picture-plane inversion sometimes dissociate perception of features and spacing in faces, and sometimes not? Toward a new theory of holistic processing. *Psychonomic Bulletin & Review 16*, 778–797.

McKone, E., Kanwisher, N., & Duchaine, B. (2006). Can generic expertise explain special processing for faces? *Trends in Cognitive Sciences, 11*(1), 8–15.

McKone, E., & Palermo, R. (2010). A strong role for nature in face recognition. *Proceedings of the National Academy of Sciences (PNAS), US, 107*(11), 4795–4796.

Milner, A. D., & Goodale, M. A. (2006). *The visual brain in action*. New York: Oxford University Press.

Mondloch, C., Le Grand, R., & Maurer, D. (2002). Configural face processing develops more slowly than featural processing. *Perception, 31*, 553–566.

Narumoto, J., Okada, T., Sadato, N., Fukui, K., & Yonekura, Y. (2001). Attention to emotion modulates fMRI activity in human right superior temporal sulcus. *Brain Research. Cognitive Brain Research, 12*(2), 225–231.

Ostrovsky, Y., Andalman, A., & Sinha, P. (2006). Vision following extended congenital blindness. *Psychological Science, 17*(12), 1009–1014.

Palermo, R., & Rhodes, G. (2002). The influence of divided attention on holistic face perception. *Cognition, 82*, 225–257.

Parvizi, J., Jacques, C., Foster, B. L., Withoft, N., Rangarajan, V., Weiner, K. S., et al. (2012). Electrical stimulation of human fusiform face-selective regions distorts face perception. *The Journal of Neuroscience, 32*(43), 14915–14920.

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.

Pascalis, O., deHaan, M., Nelson, C. A., & de Schonen, S. (1998). Long-term recognition memory for faces assessed by visual paired comparison in 3- and 6-month-old infants. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 24*(1), 249–260.

Pitcher, D., Duchaine, B., Walsh, V., Yovel, G., & Kanwisher, N. (2011). The role of lateral occipital face and object areas in the face inversion effect. *Neuropsychologia, 49*(12), 3448–3453.

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.

Purcell, D. G., & Stewart, A. L. (1988). The face detection effect: Configuration enhances perception. *Perception and Psychophysics, 43*(4), 355–366.

Rivolta, D., Palermo, R., Schmalzl, L., & Williams, M. A. (2012). An early category-specific neural response for the perception of both places and faces. *Cognitive Neuroscience, 3*(1), 45–51.

Robbins, R., & McKone, E. (2007). No face-like processing for objects-of-expertise in three behavioural tasks. *Cognition, 103*, 331–336.

Rossion, B., Gauthier, I., Tarr, M. J., Despland, P., Bruyer, R., Linotte, S., et al. (2000). The N170 occipito-temporal component is delayed and enhanced to inverted faces but not to inverted objects: An electrophysiological account of face-specific processes in the human brain. *NeuroReport, 11*, 69–74.

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

Scherf, K. S., Behrmann, M., Humphrey, K., & Luna, B. (2007). Visual category-selectivity for faces, places and objects emerges along different developmental trajectories. *Developmental Science, 10*(4), F15–F30.

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

Schmalzl, L. (2007). *Fractionating face processing in congenital prosopagnosia*. Sydney, Australia: Macquarie University.

Sengpiel, F. (2007). The critical period. *Current Biology, 17*, R742–R743.

Slater, A., Quinn, P. C., Hayes, R., & Brown, E. (2000). The role of facial orientation in newborn infants' preference for attractive faces. *Developmental Science, 3*(2), 181–185.

Tanaka, J. W. (2001). The entry point of face recognition: Evidence for face expertise. *Journal of Experimental Psychology: General, 130*(3), 534–543.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology A, 46*(2), 225–245.

Turati, C., Bulf, H., & Simion, F. (2008). Newborns' face recognition over changes in viewpoint. *Cognition, 106*, 1300–1321.

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 Science.*. doi:10.1073/pnas.1000567107.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81*, 141–145.

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception, 16*, 747–759.

Yovel, G., & Kanwisher, N. (2005). The neural basis of the behavioral face-inversion effect. *Current Biology, 15*, 2256–2262.

http://www.springer.com/978-3-642-40783-3

Prosopagnosia
When all faces look the same
Rivolta, D.
2014, XII, 95 p. 30 illus., 1 illus. in color., Hardcover
ISBN: 978-3-642-40783-3