Visual Cognition, 2013
http://dx.doi.org/10.1080/13506285.2013.772929

Routledge
Taylor & Francis Group

The composite face illusion: A whole window into our understanding of holistic face perception

Bruno Rossion

Institute of Research in Psychology and Institute of Neuroscience, University of Louvain, Louvain-la-Neuve, Belgium

Two identical top halves of a face are perceived as being different when their bottom halves belong to different faces, showing that the parts of a face cannot be perceived independently from the whole face. When this visual illusion is inserted in a matching task, observers make more mistakes and/or are slower at matching identical top face halves aligned with different bottom halves than when the bottom halves are spatially offset: The composite face effect. This composite face paradigm has been used in more than 60 studies that have provided information about the specificity and nature of perceptual integration between facial parts ("holistic face perception"), the impairment of this process in acquired prosopagnosia, its developmental course, temporal dynamics, and neural basis. Following a review of the main contributions made with the paradigm, I explain its rationale and strengths, and discuss its methodological parameters, making a number of proposals for its optimal use and refinement in order to improve our understanding of holistic face perception. Finally, I explain how this standard composite face paradigm is fundamentally different than the application to facial parts of a congruency/interference paradigm that has a long tradition in experimental psychology since Stroop (1935), and which was originally developed to measure attentional and response interference between different representations rather than perceptual integration. Moreover, a version of this congruency/interference paradigm used extensively over the past years with

Please address all correspondence to Bruno Rossion, Institute of Research in Psychology (IPSY) and Institute of Neuroscience (IoNS), Université catholique de Louvain, 10, Place du Cardinal Mercier, 1348 Louvain-la-Neuve, Belgium. E-mail: bruno.rossion@uclouvain.be

I would like to thank Jeremy Badler, Jean-Yves Baudouin, Shlomo Bentin, Thomas Busigny, Adélaïde de Heering, Giulia Dormal, Zaifeng Gao, Valérie Goffaux, Graham Hole, Suzanne Quadflieg, Aliette Lochy, Elinor McKone, Meike Ramon, Jim Tanaka, Jessica Taubert, Israr Ul Haq, Goedele Van Belle, Quoc Vuong, Andy Young, and four anonymous reviewers for their critical and helpful comments on previous versions of this paper. I am also indebted to Shlomo Bentin and Zaifeng Gao for providing me with the data of their study to reanalyse, to Lothar Spillmann for providing me with pdf reprints of his chapters during the writing of this paper and his comment on the paradoxical gap composite illusion, and to Kazunori Morikawa for his picture of the head size illusion. This work was supported by the Belgian National Fund for Scientific Research (FNRS).

composite faces lacks a baseline measure and has decisional, attentional, and stimulus confounds, making the findings of these studies impossible to interpret in terms of holistic perception. I conclude by encouraging researchers in this field to concentrate fully on the standard composite face paradigm, gaze contingency, and other behavioural measures that can help us take one of the most important challenges of visual perception research: Understanding the neural mechanisms of holistic face perception.

Keywords: Attention; Composite illusion; Congruency; Face inversion; Face perception.

PART 1: THE COMPOSITE EFFECT AND HOLISTIC FACE PERCEPTION

Introduction: The composite face illusion

For a number of years, I have been interested in understanding face perception: How does the human brain build an image, a visual representation, of a complex visual pattern such as a face? In particular, I find it fascinating that the human brain can rapidly and effortlessly extract a sufficiently detailed representation of a given face to tell it apart from other highly similar visual patterns. That is, to tell it apart from other individual faces (*individual face discrimination*). Equally interesting to me is our ability to tell that two face pictures, even of unfamiliar people, belong to the same person (*individual face matching*). In order to understand the nature of individual face perception, I have been particularly attracted to the following observation: Associating identical top halves of faces (i.e., the halves above the tip of the nose) with different bottom halves creates a compelling *visual illusion*: One cannot help perceiving the physically identical top halves as being different (Figure 1).

As with many other visual illusions, being aware that these top face halves are strictly identical does not change my perception: I am still under the persisting visual impression that the top halves are *not* the same. In the face processing literature, this visual illusion is called the *composite face illusion*.

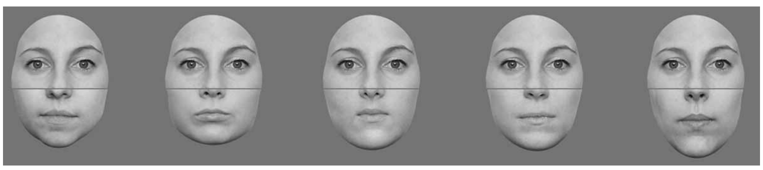

Figure 1. The composite face illusion. All 5 top halves (above the thin line) are physically identical. Yet, when they are aligned with distinct bottom halves (all of different face identities, neutral expression, taken under the same lighting conditions), they are perceived as being different.

This expression comes from the fact that *composite faces*, that is, faces in which the two halves belong to two different face identities, are used.

The composite face illusion derives from a seminal paper published 25 years ago. Andy Young and colleagues (Young, Hellawell, & Hay, 1987) aligned the top and bottom halves of celebrities' faces (e.g., the top half of Marilyn Monroe's face with the bottom half of Margaret Thatcher's face). The authors noticed that in such a composite face, the two halves fuse to form an effectively novel (unfamiliar) face (Young et al., 1987, p. 748, Fig. 1). Consequently, participants in their study found it difficult to identify familiar people from the top or bottom half of these composite faces.

Though elegant, a limitation of Young et al.'s (1987) procedure was the use of an identification task, which naturally introduces the constraint that faces are either already familiar or that they are learnt for the purpose of the study. Some years later, Hole (1994) introduced an important procedural alternative by showing that the composite face phenomenon extends to unfamiliar faces in a simultaneous matching/discrimination task. This author developed a paradigm—the *composite face matching paradigm*—in which observers take a particularly long time to match two top halves of individual unfamiliar faces when they are aligned with different bottom halves, reporting a behavioural measure of the illusion illustrated in Figure 1. To date, more than 60 published studies have followed Hole's extension of the basic method to unfamiliar face matching tasks.

The composite face effect

With the exception of Hole's (1994) study (see also Hole, George, & Dunsmore, 1999), in which faces were simultaneously presented side-by-side, the composite face paradigm is usually a *delayed* matching task of two top face halves (Figure 2). Because the bottom halves are different, observers make mistakes. That is, they tend to respond "different" for identical top

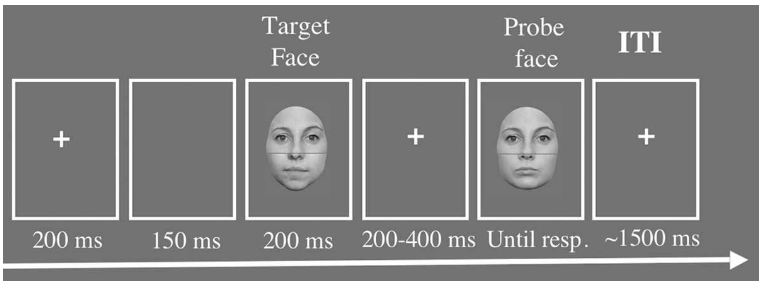

Figure 2. The composite face illusion in the context of a delayed matching task. Observers have to match the sequentially presented top halves (top = above the small gap between the face halves). The task is difficult because the top halves are erroneously perceived as being different.

facial halves, and/or they take a particularly long time to respond "same" correctly.

Crucially, it is not the mere presence of different bottom halves in the display that leads to errors, and/or to a slowing down of the response. It turns out that if the bottom halves are spatially offset from the top halves, the composite visual illusion disappears (Figure 3). Young et al. (1987) were the first to use this spatial misalignment manipulation in their original naming task of composite face celebrities, in which the parts of aligned stimuli were harder to identify than the parts of misaligned stimuli.

Consequently, the composite face matching paradigm follows the logic devised by Young et al. (1987) in that it usually has two conditions, which differ by only one factor: The spatial alignment of the bottom half relative to the top half (Figure 4). Typical observers' performance in matching the identical top halves is significantly better (i.e., more accurate and/or faster) with misaligned faces than with aligned faces.

As already mentioned, the composite face matching paradigm, which is simply referred here as the composite face paradigm, has been used in numerous studies, especially in the last decade. In writing this review paper, I have three primary intentions. The first is to create a taxonomy of the empirical work on the composite face effect, and explain how this work is fundamental for our understanding of the nature of face perception. In Part 1, although I will try to mention all published studies on the composite face paradigm, I will discuss and illustrate only a few studies in more detail, usually the ones performed by my colleagues and myself within the last decade. The review is thus selective, focusing on the studies that I know best, but the bibliography is comprehensive. The second intention is to explain as clearly as possible the rationale behind this paradigm, highlight its strengths compared to other similar experimental paradigms, explain how to use it under different circumstances, and discuss what can and cannot be inferred from it and how to improve it (in Part 2).

My third intention will be to explain why this particular composite face paradigm is fundamentally different from a "congruency" or "interference" paradigm that has been used relatively recently with composite faces. I will

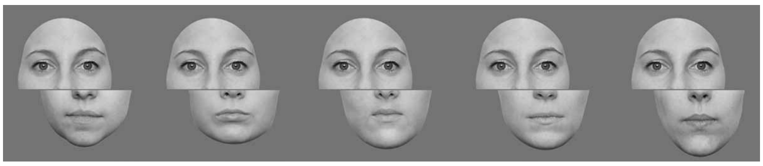

Figure 3. The composite face disillusion. All 5 top halves (above the thin line) are physically identical. If the bottom halves differ but are spatially misaligned with the top halves, one has no difficulties in perceiving the top face halves as being identical.

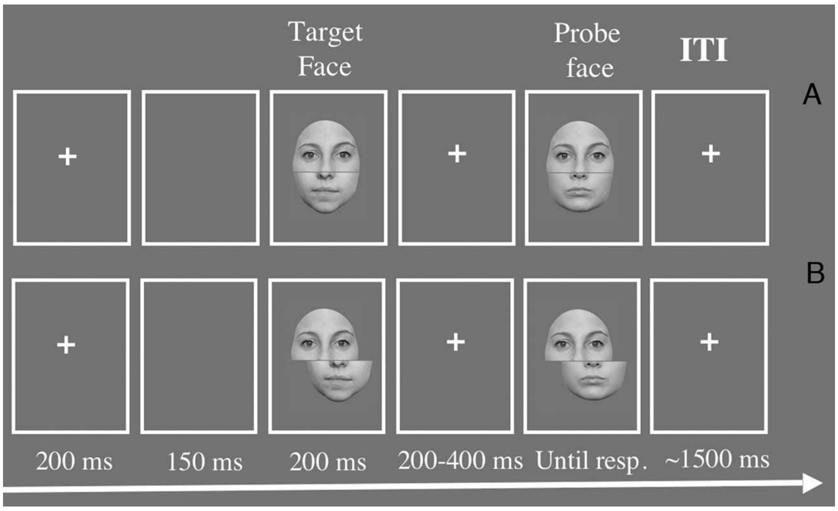

Figure 4. The composite face paradigm, in its usual context of a delayed matching task. Observers have to match the sequentially presented top halves (top = above the small gap between the two face halves). When the two face halves are aligned with each other (A), the task is difficult because the top halves are erroneously perceived as being different. (B) When the exact same stimuli are presented with their bottom halves spatially misaligned, the two top halves are readily perceived as being identical. The increase in error rates and correct response times (RTs) in the aligned face condition (A) as compared to the misaligned face condition (B) is the composite face effect.

show (in Part 3) that this congruency/interference paradigm is based on a different rationale than the standard composite face paradigm and is generally inadequate to make inferences about the specific nature of face perception. In doing this, I will dispel some myths that have arisen in the literature over the past few years. Finally, I will show that a particular version of the congruency/interference paradigm with composite faces has important built-in stimulus, attentional, and response conflict confounds. The contrast between the two approaches and paradigms will lead to the conclusion that, whereas the standard composite face paradigm measures an illusion, the congruency/interference face paradigm essentially creates the illusion of a measure.

Holistic perception of individual faces

The composite face paradigm is based on a strong visual illusion, which I take to be the clearest evidence that a human face cannot be perceived as a collection of independent parts: The perception of one part of a face is strongly influenced by the whole face. As Francis Galton (1883, p.3) once put it when describing the "human features" of a face: "One small discordance overweighs a multitude of similarities and suggests a general unlikeness."

Note that Galton here refers to a discordant part of a face that can be, or is even likely to be, the focus of attention. This point is reflected by the next sentence in this author's citation: "If any one of them (i.e., a face feature) disagrees with the recollected traits of a known face, the eye is quick at observing it, and it dwells upon the difference" (Galton, 1883, p.3).

However, what is truly remarkable about the composite face illusion/effect is that the observer attempts to concentrate, and is able to keep gaze fixation (de Heering, Rossion, Turati, & Simion, 2008) on a part (the top half of the face) that does *not* vary between the two faces. To use Galton's (1883) terminology, the observer, in fact, does *not* "dwell upon" the bottom face half. However, even if this bottom half is not fixated, its alignment with the top half nevertheless creates the perception of a whole new face (Young et al., 1987, Fig. 1).

In the composite face paradigm, when one has to compare *two* top halves, this alignment creates *two* new faces. In Galton's terms, the difference between the bottom parts of the two faces overrules the perception of the multitude of similarities between their two top parts. Therefore, this composite illusion strongly suggests that the face is perceived as a *whole*, an integrated percept. There is no way that one can perceive the top part in isolation, or before the bottom part (i.e., sequentially), to make a fast and correct decision of identity on the two top halves. This is the reason why this paradigm reflects what appears to be a fundamental aspect of face perception, in fact what may be at the heart of our special ability to recognize individual faces: *Holistic face perception*. To use Galton's (1883) terminology again, the multitude of small details of a face seem to be all perceived at a single glance.

Holistic face perception. The term "holistic" derives originally from the Gestalist view of visual perception (Koffka, 1935/1963; Kohler, 1929/1971; Wertheimer, 1925/1967; for reviews, see Kimchi, Berhmann, & Olson, 2003; Pomerantz & Kubovy, 1986; Wagemans, Elder, et al., 2012; Wagemans, Feldman, et al., 2012) that the whole is different than the sum of its parts. This term is widely used in face perception research, the human face being considered as the quintessential whole, or Gestalt (Pomerantz & Kubovy, 1986; Pomerantz et al., 2003). In line with Young et al.'s (1987) observation, in a composite face the whole is different than the sum of its parts, the whole taking properties that are novel, unpredictable, or even surprising. These elements of novelty and surprise, referred to as "emergent features" (Pomerantz & Portillo, 2011), are at the core of a Gestalt.

In the field of face processing, one generally refers to the more general term "holistic processing" rather than "holistic perception", even though it is very clear that "holistic" refers here to a perceptual phenomenon. Holistic face processing/perception has received several definitions that are not fundamentally different from each other (e.g., Farah, Wilson, Drain, & Tanaka, 1998; Maurer, Le Grand, & Mondloch, 2002; McKone, Martini, & Nakayama, 2003; Rossion, 2008; Tanaka & Farah, 1993, 2003). In line with Galton (1883), my own definition would be "the simultaneous integration of the multiple parts of a face into a single perceptual representation" (Rossion, 2008, 2009). Even though there is still a long way to go before we understand how such a process is implemented in the human brain, in terms of neural mechanisms, one finds a number of important elements in this definition. First, "holistic" means both a *process* and a *representation* of the visual stimulus. Or, to put it differently, a process that leads to, and is guided by, a representation (a representation can be defined here very loosely as a pattern of activity in the system that has a specific relationship with an external event of the physical world). Second, this representation corresponds to a visual *percept*. That is, not just a sensation but an *interpretation* of the visual stimulus based on our internal knowledge, in line with a Helmholtzian view of perception (Gregory, 1997). The fact that the composite face illusion disappears when the exact same stimulus is presented upside-down, as will be discussed later, fully supports this view. Third, it is a *unitary* percept. That is, the parts of the face are not perceived independently from the whole face. Of course, these face parts should somehow be processed independently at early stages of visual processing (for instance, in populations of neurons of the primary visual cortex that have a small receptive field). However, according to a strong holistic view of face perception, these parts would not be perceived as face-like at such early stages of processing: The first percept of the stimulus as a face would be the entire face. Finally, the parts would be integrated *simultaneously* rather than one after the other, i.e., sequentially. In summary, a face stimulus would be seen as a whole because its process is guided by an internal representation that is inherently holistic (i.e., a template). It is a template-matching process, the matching corresponding to the perception.

Admittedly, although influential authors have clearly expressed the view that the face is represented first and primarily as a whole (Sergent, 1986), and that the parts of a face would not even have a distinct representation (Tanaka & Farah, 1993, 2003), not all researchers in this field would agree with such a strong definition of holistic face perception. However, this definition does not constrain too much what will be discussed in this paper, and it will be helpful to have such a framework in mind.

Holistic/configural processing, configuration and parts: Some clarifications. In order to avoid confusion as much as possible in the remainder of the paper, it is important to make a few more points at this stage.

First, the term "holistic" refers to a *process* rather than a source of information. In my view, one could also use the terms "configural" or "configurational" to refer to this process, as in the original studies of Young et al. (1987) and Hole (1994) of the composite face effect. In fact, the terms "configural" and "holistic" were used interchangeably in the face processing literature for quite some time. It is no longer the case, with all sorts of distinctions having been introduced, and a popular view is that there are "many faces" of configural/holistic processing (Carey, 1992; Maurer et al., 2002). As I have argued elsewhere (Rossion, 2008, 2009; see also McKone & Yovel, 2009), if one targets conceptual clarity, such distinctions seem superfluous. In particular, the terms "holistic" and "configural"—when they refer to a process or to a representation—should rather be used as synonyms in face perception research.

One reason why this is not the case is that the term "configural" is ambiguous. Indeed, following Carey and Diamond (1977), researchers in the field typically distinguish between so-called "featural" and "configural" information of the face (also called first- or second-order features by Rhodes, 1988; see, for recent reviews, Bruce & Young, 2012; Bruyer, 2011; Tanaka & Gordon, 2011). A "featural" (or a part-based) piece of information is a local diagnostic cue, such as the shape of a mouth, or the colour of an eye. A "configural" piece of information refers to the position of the features/parts relative to each other (e.g., the nose above the mouth) and the metric distance between these features/parts (e.g., the interocular distance). Although this conceptual distinction between two types of information is certainly useful, the terminology used has led to important misconceptions in the field of face processing.

The first misconception is that if one has to discriminate two faces that differ in terms of "configural information" (e.g., two faces that differ in terms of interocular distance), then this would reflect "configural processing". On the other hand, if the faces differ in terms of a local feature/part (e.g., two faces that differ in terms of the shape of the nose), this would necessarily entail a "part-based" or "featural" processing (e.g., Pitcher, Walsh, Yovel, & Duchaine, 2007). This is a misconception because a change in stimulus structure does not necessarily imply a change in the way the stimulus is processed. In reality, if the face is processed holistically/configurally, a cue that is manipulated on a face stimulus is *always configural* in some sense: The perception of a local (featural) cue always depends on the other features of the face (e.g., Rhodes, Brake, & Atkinson, 1993; Sergent, 1984; Tanaka & Sengco, 1997). Thus, it is misleading to refer to certain cues only, such as relative distances between features, as being the "configural" ones.

This frequent misconception arises because the diagnosticity of the so-called "configural cues" is particularly affected by a loss of holistic/configural face perception. Indeed, a change in relative distances between features involves—by definition—several elements of the face, over a larger space, than a featural/local cue. It follows that if holistic/configural perception is impaired, for instance following inversion (see later) or acquired prosopagnosia (e.g., Ramon, Busigny, & Rossion, 2010), then the relative distances can be more difficult to perceive than the featural/local cues (e.g., Barton, Press, Keenan, & O'Connor, 2002; Freire, Lee, & Symons, 2000; Goffaux & Rossion, 2007; Rhodes et al., 1993; Sekunova & Barton, 2008; for a full discussion of this issue, see Rossion, 2008, 2009).

To avoid such misunderstandings, I believe that when referring to a *process* or to a *representation*, the field of face perception would gain a lot of clarity by using the term "configural" only as a synonym of "holistic", while keeping the distinction between local features and relative distances. In this paper, I will not refer much to the term "configuration" or "configural cues". Unless specified, the term "configural" will rather be used to refer to a process/representation and thus indeed as a synonym of "holistic" throughout.

A second important point is the following. When I write that the perception of a local face part always depends on the other parts, I do not mean only the parts that are available in the physical stimulus but also the parts of the holistic template that help in perceiving the face. Indeed, the face could be partially occluded, for instance. Or, a small part only, such as the eyes and eyebrows, could be available. Yet, for a typical observer, holistic processing *can* be applied to such a partial face stimulus. In other words, measuring holistic face perception does not mean that the whole stimulus needs to be physically present. This point needs to be stated to avoid misconceptions such as the criticism of the holistic account of face inversion on the basis of inversion effects found for a subset of parts of the face (Leder & Bruce, 2000; Rakover, in press; Rakover & Teucher, 1997; but see Bartlett, Searcy, & Abdi, 2003).

A third point refers to the fact that, according to a holistic view of face perception, the parts of a face would be processed *simultaneously* ("at a single glance"), rather than sequentially. This view does not imply that the parts have the same weight in face perception: Some parts are certainly more diagnostic than others when perceiving faces, for instance the region of the eyes (e.g., Davies, Ellis, & Shepherd, 1977; Gosselin & Schyns, 2001; Haig, 1985, 1986; Sheperd, Davies, & Ellis, 1981). The relative saliency of parts may vary according to the observers' gaze fixation and the task at hand (e.g., Gosselin & Schyns, 2001; Smith, Cottrell, Gosselin, & Schyns, 2005). It follows that if one presents isolated parts to the system—whether they are arbitrarily (e.g., Leder & Bruce, 2000; Rakover, in press; Rakover & Teucher,

1997) or randomly (Gosselin & Schyns, 2001; Haig, 1985) defined—a holistic face representation may be triggered faster, or more strongly, by certain parts of a face (e.g., the eyes) than by others. Such an observation should not necessarily be interpreted as evidence for sequential processing of facial parts (another misconception, see, e.g., Schyns, Jentzsch, Johnson, Schweinberger, & Gosselin, 2003; Smith, Gosselin, & Schyns, 2007).

A fourth point refers to the distinction between holistic and analytic processing. As much as I do not want to distinguish between a "holistic" and "configural" mode of processing, it makes perfect sense to me to distinguish between what we have defined as holistic face processing and its opposite: The processing of a stimulus sequentially, local part by local part. Although this is not a very efficient way of processing a face, this analytical processing mode *can* be used to individualize faces. However, contrary to holistic processing, there are not many reasons to believe that this analytical mode of processing is particularly interesting if one wants to understand what is specific about the nature of face perception (see earlier).

Fifth, although I discussed this issue previously (Rossion, 2008, 2009), I would like to stress again that the emphasis on holistic/configural face processing does not at all mean that facial parts are not important to recognize individual faces. Unfortunately, this is also a frequent misconception in the field (e.g., Cabeza & Kato, 2000). Of course, facial parts *are* important, a point that was precisely emphasized by Young et al. (1987) at the end of their seminal paper. Facial parts are the building blocks of our ability to individualize faces. The holistic view simply states that a facial part is not perceived independently of the other parts: The parts are necessarily grouped into a holistic representation. For this reason, if anything, the role of facial parts is even more important in a holistic processing framework than in an analytical processing framework: In a holistic processing framework, modifying a face part changes the whole face. Again, for this reason, when observers match faces that differ by local parts, it does not mean that they perform part-based processing.

Finally, it is important to state that even though holistic face processing refers to a single representation in the present framework, this process can take place at different degrees of resolution. Justine Sergent (1986) expressed this view very well, in a remarkable theoretical paper. That is, a coarse holistic representation may be sufficient to detect a face in a visual display, but not to individualize it. In order to individualize a face, one needs to build a holistic percept that is detailed enough to be able to distinguish it from percepts built from different faces. This is the difference between holistic face perception as measured when one has to detect "a" face in a visual display that has no visible face parts (a Mooney face, Mooney, 1957; Moore & Cavanagh, 1998; or a Arcimboldo painting, Hulten, 1987; see Figure 5A), and holistic face perception as measured in the composite paradigm, when one has to match *individual* faces. If holistic face perception can be present at different degrees of resolution, or spatial scales, it may be that several holistic representations are necessary to process a face. Alternatively, in a more dynamic and integrated framework, one could envision a coarse-to-fine process in which an originally coarse holistic face percept is gradually refined in order to individualize the face (Figure 5B). This point is very important because we are concerned here primarily with the holistic perception of *individual* faces.

Inversion. With these conceptual clarifications in hand, let me come back to composite faces. Another major reason why the composite face illusion never fails to fascinate me is that it disappears when faces are presented upside-down (Figure 6). Again, in their seminal study using a famous face identification task, Young and colleagues were the first to use inversion with composite faces and these authors showed that the parts of aligned inverted stimuli were easier to identify than the parts of upright stimuli (Young et al., 1987, Exp. 2). And, rather than comparing aligned and misaligned faces as in the vast majority of subsequent studies, Hole (1994) actually introduced the composite face matching paradigm by showing a better matching performance (faster RTs) for inverted than upright parts in composite faces.

Figure 5. (A) In binarized stimuli (i.e., with pixels being either white or black), the parts of faces (the two stimuli on the left) are not perceived as face-like if one cannot use the whole face configuration (the same stimuli on the right, with the 4 parts grouped together). (B) In a coarse-to-fine perceptual process, the initial representation of a face is that of the whole face, not of separated face parts. A face can already be detected from an initial representation such as the one on the left, but this representation is too coarse to individualize the face. Following a refinement of the face representation over time, it can be individualized. Importantly, in such a dynamic coarse-to-fine mode of processing, the parts are never represented as face-like independently of the whole face: It is a single holistic process. Both face detection (or categorization of stimulus as a face) and face individualization depend on the same holistic representation that evolves dynamically over time (figure adapted from Sergent, 1986). To view this figure in colour, please see the online issue of the Journal.

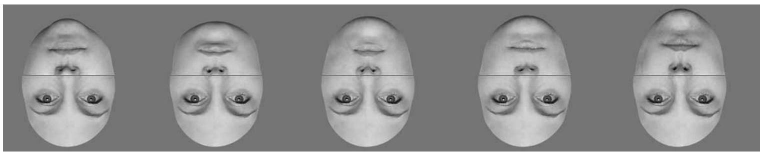

Figure 6. The inverted composite face disillusion. This is the same figure as Figure 1, but the faces have been vertically flipped. All 5 "top" halves (here at the bottom of the display, below the thin line) are physically identical and, unlike at upright orientation (Figure 1), they are no longer perceived as being different.

These observations nicely demonstrate that turning a stimulus upside-down does not merely make face processing more difficult. That is, inversion is not a manipulation that affects face processing merely *quantitatively* (Sekuler, Gaspar, Gold, & Bennett, 2004; Valentine & Bruce, 1998). Instead, it seems that part-based, analytical processing is relatively well preserved for inverted faces, but that the perception of the individual face as a whole is impaired by inversion. These observations fully support a *qualitative* view of face inversion (Rossion, 2008). I have recently attempted to explain what happens when the face is upside-down in terms of a reduction, or shrinking, of the *perceptual field* (Rossion, 2009; see Figure 7). This concept, which refers to *the area of vision where the observer can extract diagnostic visual information for the task*, is close to notions such as the *functional field of vision*, the *perceptual span*, the *visual span*, or the *span of effective vision*, as defined initially and used mainly in the reading literature (Rayner, 1975, 1998; see also Jacobs, 1986; Reingold, Charness, Pomplun, & Stampe, 2001).[1] According to the perceptual field hypothesis, when fixating a specific part of an upright face, the right eye for instance, one would perceive the whole face—thanks to the matching with a holistic template. However, when the face is upside-down, one would perceive only the right eye (Figure 7).

Recently, we extended the gaze-contingent window technique (McConkie & Rayner, 1975) to face perception (Van Belle, de Graef, Verfaillie, Rossion, & Lefèvre, 2010; see Figure 8). By restricting the field of vision online to roughly one face part, we decreased the face inversion effect. In contrast, when we gaze-contingently masked the fixated part, thus promoting holistic processing, the face inversion effect increased (Figure 8). This observation provides empirical support for the perceptual field account of the face inversion effect.

[1] Jung and Spillmann (1970) introduced earlier the notion of the *perceptive field* (see Spillmann, Ransom-Hogg, & Oehler, 1987). However, these authors meant the receptive field as determined by psychophysics (e.g., Neri & Levi, 2006), as opposed to the receptive field determined in neurophysiology at the level of the single neuron.

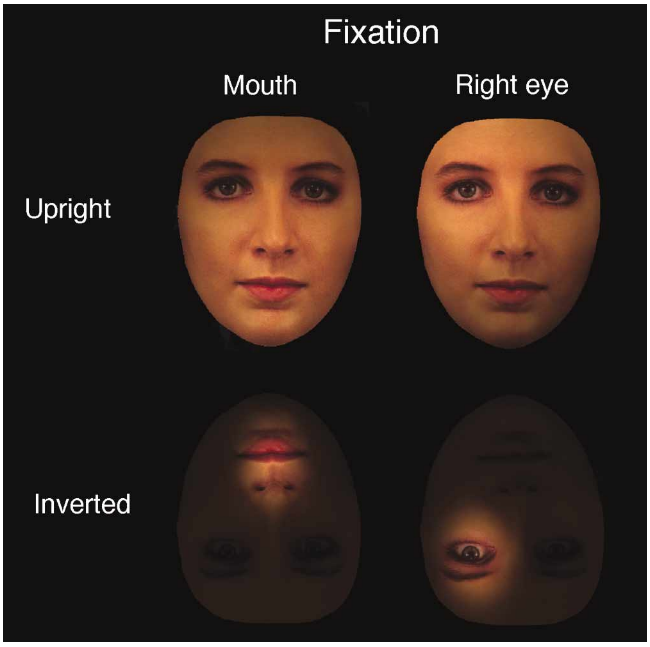

Figure 7. Because of holistic perception, fixating a part does not prevent us from seeing the whole face when it is presented at upright orientation. In contrast, when the face is inverted, perception cannot be guided by an internal holistic face representation, so that the perceptual field would be reduced and fixating a part would mean perceiving only that part at a fine-grained level of detail (figure adapted from Rossion, 2009). To view this figure in colour, please see the online issue of the Journal.

If the perceptual field is reduced when faces are presented upside-down, observers' performance in matching identical "top" halves of composite faces can no longer be influenced by their bottom halves. Consequently, observers' performance in the composite face paradigm improves (e.g., Hole, 1994). As noted by Young et al. (1987), this improvement is paradoxical: For once, people may perform better with faces cut in half, or inverted, than with whole upright faces!

How are faces special(ly) holistic?. The fact that the composite face effect disappears/decreases with inversion shows that the effect does not merely reflect a general process, i.e., one that would be applicable to *any* visual shape. Of course, nonface object shapes are also perceived holistically or configurally, their parts being integrated into wholes. However, an upright face appears to represent the ultimate form of a Gestalt, its parts being particularly strongly interdependent with each other. Accordingly, there is a

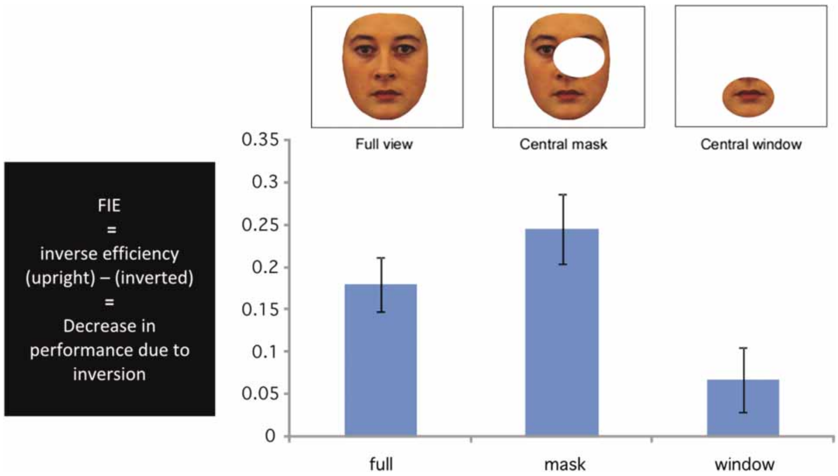

Figure 8. Gaze-contingency and face inversion (Van Belle, de Graef, Verfaillie, Busigny, & Rossion, 2010). The magnitude of the face inversion effect in a 2-alternative forced choice individual matching task is decreased when observers have to process the face part by part, due to the restriction of their field of view to a small gaze-contingent window. In contrast, when the central window of fixation is masked, promoting holistic processing, the inversion effect increases. To view this figure in colour, please see the online issue of the Journal.

general consensus that faces are perceived *more* holistically than other objects (Biederman & Kalocsai, 1997; Tanaka & Farah, 1993, 2003).

What does this mean exactly? One can only speculate here, as the general issue of how individual parts combine into perceptual wholes remains a central challenge for visual perception theorists. In my opinion, there are two important differences between holistic/configural perception of faces and holistic/configural perception of nonface objects, which makes a comparison between the two very difficult.

First, compared to nonface object shapes, faces are extensively experienced and they all share the same basic structure (symmetry, round shape, two eyes on top of central nose and mouth, etc.). These characteristics favour the construction and use of a template, or a "schema" (Goldstein & Chance, 1980), to perceive faces. This could be the reason why faces, compared to other objects, are more easily detected in visual displays that contain little part information, such as binary "Mooney" images (Moore & Cavanagh, 1998; see Figure 5A). For the same reason, if only part of a face is presented to the visual system (the eyes region for instance), a whole face representation might be activated automatically. This makes it very difficult or even impossible to assess whether the representation of a whole face is truly different than the summed representation of its parts (i.e., a nonlinearity) by contrasting the behavioural or neural response to a part versus the whole face (e.g., Freiwald, Tsao, & Livingstone, 2009; Gold, Mundy, & Tjan, 2012; Kobatake & Tanaka, 1994). Also, this makes it difficult to use facial parts to study grouping processes the same way as one can use object parts to study holistic processing of nonface object shapes (e.g., Pomerantz et al., 2003).

The second difference between holistic/configural perception of faces and holistic/configural perception of nonface objects is that holistic processing does not only help face detection (McKone, 2004; Rossion, Dricot, Goebel, & Busigny, 2011; Taubert, Apthorp, Aagten-Murphy, & Alais, 2011), but also face *individualization*. In particular, with the composite effect, we deal with this second level: The individualization of the stimulus, a process that requires a more detailed representation than what is needed for face detection. At this individual level, it is unclear whether nonface objects are processed holistically: The individualization of a nonface object from another member of the same category appears to rely essentially on part-based analysis (Biederman & Kalocsai, 1997; Farah, Klein, & Levinson, 1995). In this perspective, what is truly special about faces as compared to other complex visual object categories is not that faces are processed holistically, *or* that they can be processed at a fine-grained level of resolution. What is special is that faces are processed holistically *at* a sufficiently fine-grained level of resolution to individualize members of the face class (see Busigny, Joubert, Felician, Ceccaldi, & Rossion, 2010). In line with this claim, studies that have applied the composite paradigm with nonface objects have failed to report any composite effect (dog pictures, Robbins & McKone, 2007; car pictures, Macchi Cassia, Picozzi, Kuefner, Bricolo, & Turati, 2009; novel objects called "Greebles", Gauthier, Williams, Tarr, & Tanaka, 1998, or "sticks", Taubert, 2009).

The role of a template derived from visual experience. When discussing the issue of face inversion, I made the point that holistic perception depends heavily on an internal representation. In line with this, studies with typical observers have shown that the magnitude of the composite face effect is increased for categories of human faces whose morphological type is the most frequently experienced. For instance, "same-race faces" lead to a larger composite face effect than "other-race faces" (Michel, Rossion, Han, Chung, & Caldara, 2006; see also, for modulation of this effect for "racially ambiguous" faces by "race" categorization, Michel, Corneille, & Rossion, 2007, 2010). The same phenomenon has been reported for same-age faces (Susilo, Crookes, McKone, & Turner, 2009). Moreover, extensive visual experience, even at adulthood, with a specific regime of faces—for instance children's faces for schoolteachers—increases the composite face effect for such faces (de Heering & Rossion, 2008; see also Kuefner, Macchi Cassia, Viscovo, & Picozzi, 2010; see Figure 9).

Collectively, these studies suggest that even in typical observers, and within the face domain, the integration of individual facial parts into a whole is constrained by our long-term visual experience, and thus by our internal representations of the visual world.

Supporting this suggestion, the composite face effect does not disappear linearly with the angle of plane rotation from the upright face (0°). Rather, it is equally large for stimuli presented at 0° until 60° rotation, then decreases abruptly at 90° and remains stable until complete inversion of the stimulus (see Figure 10, from Rossion & Boremanse, 2008[2]). This observation is particularly interesting because it shows that holistic perception might be at play only for faces that we experience in real life. Once the face has reached an orientation that is almost never experienced (≥90°), holistic processing is absent, or at least strongly reduced, and does not decrease further. This nonlinearity suggests that misoriented faces are not first realigned by means of linear rotation processes that would work independently of internal representations

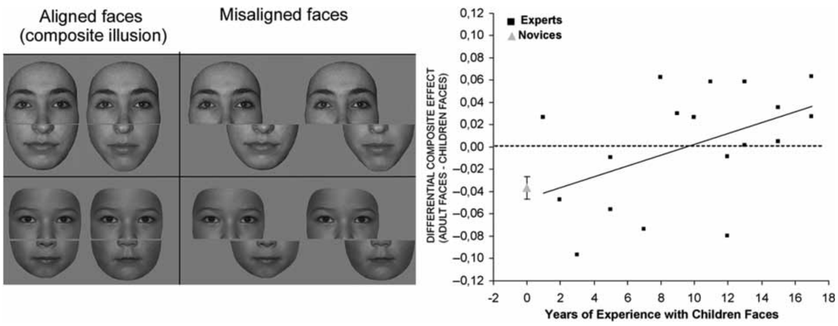

Figure 9. In the study of de Heering and Rossion (2008), the composite face effect was measured for adult and child faces, in adults that either had limited experience with children's faces (novices) or extensive experience with such faces (schoolteachers). The differential magnitude of the composite face effect (adults vs. children faces), as measured in correct RTs, was larger for experts than novices, and was positively correlated with the number of years of experience with children's faces (since the effect is measured in RTs, the effect reflects the subtraction of RTs for children faces from the RTs for adult faces, a negative value meaning a larger effect for adult than children faces).

[2] In a similar study, Mondloch and Maurer (2008) also observed that the composite effect was no longer significant beyond 90°, and that there was no difference between further orientations. Thus, although the authors concluded that the composite effect decreases linearly with rotation, it seems that their data (see their Figure 2a) are rather in agreement with the findings displayed here in Figure 10. Additionally, it is also possible that the pattern of RT data of their participants, which was not reported, would have been compatible with the conclusion of Rossion and Boremanse (2008).

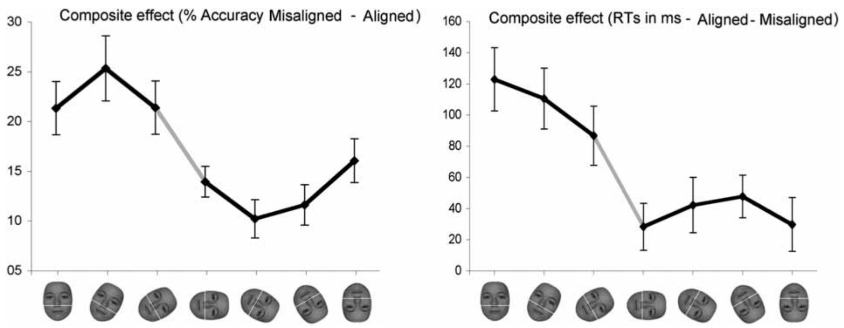

Figure 10. For both accuracy rates and correct RTs, there is a large and abrupt drop of the composite face effect between 60° and 90°, with no further decrease after this orientation. Adapted from Rossion and Boremanse (2008).

derived from visual experience. Rather, it seems that holistic face perception depends indeed on an upright, experience-derived, face template.

Despite the role of visual experience, a large composite face effect is found as early as 4 years old (de Heering, Houthuys, & Rossion, 2007; Macchi Cassia et al., 2009; see also Carey & Diamond, 1994, for evidence in 6- to 10-year-old children with a naming task; Mondloch, Pathman, Maurer, Le Grand, & de Schonen, 2007, in 6-year-olds; Susilo et al., 2009, for 8- to 13-year-old children) and it has even been reported in 3-month-old children—but not newborns—using an adaptation paradigm with eye movement recordings (Turati, Di Giorgio, Bardi, & Simion, 2010). Thus, it seems that a relatively limited visual experience can be sufficient to tune the system to process faces holistically.

The nature of the holistic face representation. What is the nature of the holistic face representation? Goffaux and Rossion (2006) attempted to answer this question by manipulating spatial frequency information in composite faces. Spatial frequencies (SF) refer to the various resolution ranges composing an image, with low SF (LSF) depicting the coarse structure of the image (e.g., the coarse shading of a face) and higher SF (HSF) representing the finer details of the image (e.g., eyelashes, skin texture, etc.) (Morrisson & Schyns, 2001) (see Figure 5). According to a long-standing hypothesis, holistic face perception is supported relatively more by low- as opposed to high-spatial frequencies (LSF vs. HSF; Sergent, 1986). In principle, this view is trivial: Holistic perception is defined as the integration of facial parts over the whole face, and it makes sense that one needs to rely essentially on variations of luminance at large scales to integrate parts that are spatially distant. However, even without considering the issue of the difference in contrast provided by LSF and HSF, this hypothesis is more difficult to test than it seems, and is almost as complicated as clarifying the importance of certain SF bands in face processing in general (i.e., independently of holistic perception) (Sergent, 1986). One reason for this difficulty is that holistic perception does not require the whole face to be in play. Two parts of a face that are close to each other, such as an eye and its eyebrow, will also benefit from the capacity of the system to process the parts holistically, i.e., as an integrated unit. Hence, holistic perception is certainly useful to extract information at multiple spatial scales, including small ones. A second issue is that holistic perception may take place at different levels (detection and individualization), requiring different degrees of resolution. A third issue is that filtering out the high spatial frequency content, for instance by applying a relatively severe low-spatial frequency cutoff (<8 cycles/image) makes the configuration of the face its main property: Local features are not available. Conversely, filtering out the low-spatial frequencies of a face does not prevent the visual system from generating low spatial frequencies from such a high-pass filtered stimulus (Ginsburg, 1978; Sergent, 1986). It follows that low spatial frequencies are, at least partly, included in *any* representation of a face in the brain.

Given these issues, comparing LSF and HSF filtered faces directly may not lead to any advantage for LSF faces, unless one uses a paradigm in which the presence of HSF information provides additional cues that can be detrimental for performance. Interestingly, this is precisely the case for the composite face paradigm. To perform the task well (i.e., reduce the integration of the bottom with the top) on aligned stimuli, it is useful that detailed information on the fixated top part is available. Having HSF information available on the top half is likely to help participants perform the task better than when only LSF information is available. In addition, because the bottom half is not fixated, most of the disrupting information coming from the bottom half will be provided by LSF. Because of these two factors combined, it is easy to predict that the composite face effect should be larger for LSF than HSF faces. Goffaux and Rossion (2006) tested this hypothesis and found indeed an increased composite effect for LSF faces, and a decreased composite effect for HSF faces. In a replication, there was also a larger effect for LSF than middle SF (8–32 cycles/image) (Figure 11). These findings were taken as evidence for a dominance of LSF in holistic face perception, even when having to individualize the face.

This effect is in line with the idea that three-dimensional shape rather than surface-based information (e.g., colour, texture) supports holistic perception. Indeed, contrast-reversed faces, in which surface cues are no longer diagnostic, are associated with a composite face illusion and substantial composite face effects (Hole et al., 1999; Taubert & Alais, 2011; see Figure 12). More recently, we found little if any contribution of surface-based cues as compared to shape in generating the composite face effect (Jiang, Blanz,

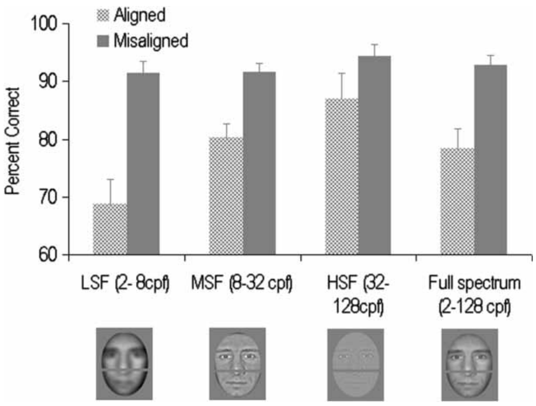

Figure 11. The composite face effect is larger for faces whose high spatial frequencies have been filtered out (LSF faces). Figure adapted from Goffaux and Rossion (2006), Exp. 4).

& Rossion, 2011; see Figure 13). The reasons for this effect have been fully discussed in Jiang et al. (2011). In brief, when faces differ by shape information, the shape of the global contour of the face and the relative size of the head are highly salient cues for individual face matching/discrimination. Moreover, "configural information", or relative distances between internal features of the face (e.g., mouth–nose, interocular distance, etc.) can also be diagnostic. In contrast, the surface cues that are diagnostic for face individualization either have to be resolved locally (e.g., colour of the lips), or *can* be resolved locally (e.g., darkness of eyebrows, colour of the eyes, skin colour).

Holistic face perception is functional. In the composite face paradigm, because the two bottom halves differ, they create the perception of two different whole faces. Therefore, in the context of a matching task, performance decreases at judging whether two top faces are the same in the aligned as compared to the misaligned condition. In this particular

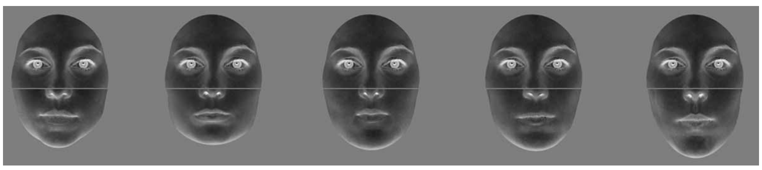

Figure 12. The composite face illusion with contrast-reversed faces. All 5 top halves (above the thin line) are physically identical, and in fact are the same faces as presented in Figure 1. Despite contrast reversal, the top halve faces are perceived as slightly different due to their alignment with distinct bottom halves. Even though the illusion is not as compelling as with typical faces, no study so far has reported a significant decrease of the composite face effect with contrast-reversed faces.

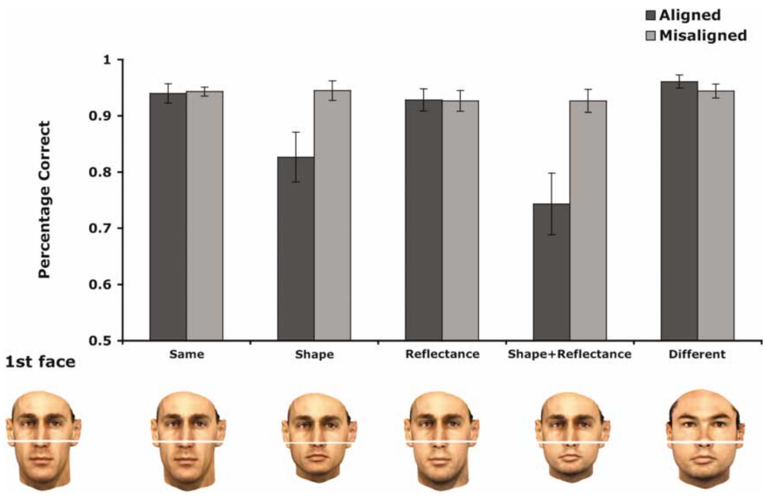

Figure 13. Figure adapted from the study of Jiang et al. (2011), showing that the composite face effect is largely accounted for by variations in the shape of the face rather than in surface cues (also called reflectance or pigmentation, that is, texture and colour). To view this figure in colour, please see the online issue of the Journal.

context, holistic perception is thus measured as a *negative* effect on performance. However, outside of that context, the ability to see an individual face as a whole (i.e., to perceive all parts in a single unified representation) should provide a substantial advantage for the observer, making it easier and faster to recognize a previously seen face and categorizing other faces as being novel. Thanks to this holistic process, the observer does not have to check every single part of a face in turn to make such judgements. Thus, as I have argued previously (Rossion, 2009), holistic face perception is certainly *functional*: If holistic face perception is impaired, face recognition performance should be impaired. Evidence supporting this claim comes from several sources.

Inversion (again). In normal observers, picture-plane inversion, which dramatically decreases face recognition performance (Hochberg & Galper, 1967; Yin, 1969), is also associated with a loss of holistic face perception as measured with the composite face paradigm and other paradigms (e.g., Sergent, 1984; Tanaka & Farah, 1993; Young et al., 1987; see, for reviews, Rossion, 2008, 2009). In the same vein, though less well documented, inverted faces of a nonexperienced morphology, such as "other-race" or "other-age" faces, are also both less well recognized (e.g., Malpass & Kravitz, 1969; Meissner & Brigham, 2001) and processed less holistically (e.g., Michel, Rossion, et al., 2006; Tanaka, Kiefer, & Bukach, 2004; see, for a review, Rossion & Michel, 2011).

Acquired prosopagnosia. Patients presenting with acquired prosopagnosia—typically the impairment in face recognition following brain damage (Bodamer, 1947; Ellis & Florence, 1990)—present with a strongly reduced or even abolished composite face effect (Busigny et al., 2010; Ramon et al., 2010). This impairment is in line with impairment in holistic face perception, as measured by a variety of other indexes (for a review, see Ramon et al., 2010). Moreover, gaze contingency shows that these patients' perception of isolated facial parts is preserved, or relatively less affected, than their perception of whole faces (Van Belle et al., 2011; Van Belle, de Graef, Verfaillie, Rossion, & Lefèvre, 2010), suggesting that their impairment in holistic face perception is a critical marker of their prosopagnosia.

Long-term impairments in face recognition. Individuals deprived of patterned visual input by bilateral congenital cataracts for 3–6 months after birth and having difficulties in face recognition may also display a reduced composite effect (Le Grand, Mondloch, Maurer, & Brent, 2004), although they may recover it after years of experience (de Heering & Maurer, in press). Some studies have also found that holistic perception as assessed by various measures (whole–part, inversion, and composite effect) is weaker in cases of congenital prosopagnosia (Avidan, Tanzer, & Berhmann, 2011; Palermo et al., 2011; but see Le Grand et al., 2006; Schmalzl, Palermo, & Coltheart, 2008).

Gaze contingency. Finally, and most directly, when perception is limited to roughly one face part at a time through gaze contingency, the performance of a normal observer may decrease to almost the level of a patient with prosopagnosia (Van Belle et al., 2011; Van Belle, de Graef, Verfaillie, Rossion, & Lefèvre, 2010).

Considering all these observations, holistic perception—as measured by the composite face effect—seems to be indeed important, or even necessary, for efficient face recognition.

Correlating holistic face perception and face recognition performance. Despite the evidence reviewed earlier, the value of the composite face paradigm has recently been challenged because the effect measured with this paradigm is not correlated with face recognition performance across individuals (Konar, Bennett, & Sekuler, 2010). Here, I would like to take that opportunity to discuss the general issue of a correlation between measures of holistic face perception and face recognition performance, and what can and cannot be inferred from such a correlation (or lack thereof).

Variability in face recognition performance and the rationale for correlation measures. In recent years, the view that we are all experts at face recognition, a view that largely dominated the field for decades (Carey, 1992)—has been challenged, with many behavioural studies showing that humans' face recognition abilities in fact vary tremendously (e.g., Germine, Duchaine, & Nakyama, 2011; Megreya & Burton, 2006; Russell, Duchaine, & Nakayama, 2009; Wilmer et al., 2010). Some people appear to be very good at face recognition ("physionomists" or "super-recognizers"; Russell et al., 2009), and some people appear to be quite poor at it. Most people seem to present with face recognition performances around the average, as though face recognition performance obeys a Gaussian distribution in the normal population (Bowles et al., 2009). Whether the poorest performers merely represent the lower tail of a normal distribution of face recognition ability, or should be defined as cases of developmental/congenital prosopagnosia (Behrmann & Avidan, 2005; Duchaine & Nakayama, 2006a), will have to be determined. Irrespective of the answer to this question, human variability in face recognition ability has prompted an increasing number of recent studies in this field to use correlation measures to understand this function.

In particular, one may be interested to know whether face recognition ability is correlated with holistic face perception. At first glance, this is a reasonable objective, and if the answer to this question is positive, it could reinforce the functional link between holistic perception and face recognition, and thus the importance of the former for the latter (Wilmer, 2008). However, this objective may be difficult to reach, for obvious reasons.

First, it is worth reminding that irrespective of the quality of the test used (for instance the widely used Cambridge Face Memory Test, CFMT; Duchaine & Nakayama, 2006b), face recognition performance as assessed in a given test will vary across individuals due to many general factors (perceptual, attentional, memory, motivational, decisional, etc.) that have nothing to do with the recognition of faces per se. If one aims at measuring face recognition ability in a given individual, these factors should be neutralized as much as possible.

Second, many of these factors will also influence a behavioural measure of holistic perception. When using the composite face paradigm in two conditions of interest (e.g., normal vs. contrast-reversed faces) at the group level, the variance due to these factors (i.e., the noise) cancels out as the size of the group increases, providing a chance to obtain a difference between the conditions. However, a measure of holistic perception in a single individual tested with normal faces can only be affected by many general factors, especially if it is based on a relatively small amount of trials. Some of these factors vary a lot across individuals, and some of these factors can vary also substantially within the same individual from session to session of recording. Thus, it is not surprising that despite the quality of the composite face paradigm, the correlation across individuals of the exact same measure of holistic perception tested twice in this paradigm is limited (e.g., split-half reliability = .52 in Laguesse & Rossion, 2011; .43 in Wang, Li, Fang, Tian, & Liu, 2012; .65 in Zhu et al., 2010), and part of this correlation is also certainly driven by general factors, not just holistic face perception.

If two measures taken twice in the exact same test are not highly correlated across typical individuals, one should not expect a high correlation between one of these measures and another measure like face recognition performance taken in the same individuals. Also, when individual data is displayed (Avidan et al., 2011; Ramon et al., 2010), some normal observers may not show any significant composite effect, neither in accuracy rates nor in RTs. Does it mean that these observers do not perceive faces holistically? Certainly not. It could equally mean that, as in any experiment, the manipulation did not work in that single recording for that participant because there was too much noise in his/her data (i.e., undesirable factors affecting behavioural performance). Moreover, a participant who emphasizes accuracy in the task may sometimes make very few mistakes, even in the critical condition (same aligned trials with different bottom face halves). However, he/she will usually be slowed down for trials in that condition. Therefore, his/her effect will be reflected in RTs; in other observers it will rather be reflected in accuracy rates. Interindividual variance may thus be distributed in the two variables, which should probably be combined to obtain a better approximation of the magnitude of holistic perception in the experiment.

In summary, a behavioural measure obtained by the composite effect is a noisy approximation of a given process (holistic face perception), divided into two dependent variables. It is not surprising that it is not very well correlated with another noisy approximation of a given process (face recognition performance), especially if each measure is the outcome of only a single test, with a limited amount of trials. Even more so when, like Konar et al. (2010), both the measures of holistic perception and face recognition performance are obtained in individuals from a single test, general factors affecting performance are not neutralized, and correlations are computed only separately for RTs and accuracy rates. In short, contrary to Konar et al.'s claim, a weak, or even an absence of correlation cannot be taken as casting doubt on the importance of holistic perception for face recognition.

(Weak) correlations can be found in the composite face paradigm. A recent study (Wang et al., 2012) replicated the absence of correlation by Konar et al. (2010) but then isolated the face-specificity measure of recognition performance by subtracting performance at recognizing nonface objects. The composite face effect in correct RTs correlated significantly with this face-specific measure. Nevertheless, despite the large number of participants (>300), the correlation remained weak ($r = .13$), suggesting that holistic perception (and face recognition performance) cannot be captured in a single

measure, and that many factors contribute to the behavioural performance in this task. Another study (Avidan et al., 2011) took advantage of the increased variance between individuals with poor face recognition ability ("congenital prosopagnosia") and found that the composite effect, in RTs, correlated ($r = .61–.72$) with the abnormality of performance on diagnostic face processing tasks (but see de Heering & Maurer, in press).

These correlations are interesting, and serve to corroborate the important role of holistic perception in face recognition. Nevertheless, one should not conclude from small correlations such as the one found by Wang et al. (2012) that holistic face perception is only weakly important for face recognition performance. It may be weakly related, but *critically* related: If a face cannot be perceived holistically, face recognition performance can be massively impaired, as we observed in cases of acquired prosopagnosia. To put it differently, holistic perception may be *necessary* for face recognition, but not *sufficient*.

Finally, researchers generally assume that there is a certain degree of individual variability in the magnitude of holistic face perception, i.e., that some people are strong holistic face perceivers, whereas others are weak holistic perceivers. This is not necessarily true. It may well be that holistic perception is a necessary entry step for processing faces efficiently and that it varies very little across individuals. Rather, interindividual variance in the behavioural composite face effect could be due to other factors, as suggested by the relatively limited split-half reliability of the task. To conclude, measuring holistic face perception in individuals certainly requires more sensitive approaches, and approaches that can isolate the perceptual process from general sensory, mnesic, attentional, and decisional/response output factors. The next sections will address this latter issue.

Neurofunctional locus

Capturing a perceptual phenomenon in neuroimaging. The composite face illusion is a perceptual phenomenon, and the studies reviewed earlier illustrate very well an approach that emphasizes phenomenology as a method for describing phenomena and collecting data. This approach has its roots in Gestalt Psychology and its scientific aim is to discover and describe structural laws of visual experience by the systematic and controlled variation of a phenomenon by independent variables (Sinico, 2003). Nonetheless, as with other perceptual phenomena, it is essential that we understand the *neural mechanisms* of the composite face illusion, because they could be fundamental to resolve one of the greatest puzzles of visual perception: How the brain can integrate different parts of a visual stimulus to form a whole configuration. Understanding this issue for faces—holistic face perception in terms of brain mechanisms—will go a long way towards understanding high-level vision, and for this reason we need to reconcile phenomenology with a neurophysiological approach of visual perception (Spillmann, 1999).

I suspect that this objective will not be reached by "simply" measuring the response of single neurons to face parts versus whole faces (e.g., Freiwald et al., 2009; Kobatake & Tanaka, 1994). That is, one needs to consider a more global level of organization in the system, and how *populations* of neurons may code for such holistic representations (i.e., a holistic approach to Gestalt perception at the neural level; Spillmann & Ehrenstein, 1996). In humans, the initial approach that we took was to use the composite face illusion and insert it into a face-identity adaptation paradigm in functional magnetic resonance imaging (fMRI), a method that indirectly measures the response of populations of neurons at a spatial resolution of several mm^3. Upon repeated presentations of the same individual face, some areas of the brain will show reduced fMRI activity as compared to when different individual faces are presented in succession (Grill-Spector & Malach, 2001). Such a release from neural adaptation, habituation, or repetition suppression effect (Grill-Spector, Henson, & Martin, 2006) can only be found if the areas are sensitive to the differences between individual faces. Based on this, we presented blocks of identical top halves of faces (one face every 1500 ms), asking participants to detect a slight change of colour on some of these top halves. This task was used to ensure that they focused on the top face halves and performed equally well in all conditions. In the condition of interest, bottom halves were from different faces, leading to the visual impression of a succession of different top halves in the paradigm (as in Figure 1). Because of this perceptual illusion of different whole faces, face-sensitive areas of the visual cortex showed a release from adaptation, as compared to conditions in which the bottom halves were identical (Figure 14; Schiltz & Rossion, 2006). This was especially true in (but not only) a small face-selective area of the right middle fusiform gyrus, termed the "fusiform face area" ("FFA"; Kanwisher, McDermott, & Chun, 1997). Importantly, this release from adaptation was not found when the two halves were spatially misaligned (Figure 14).

In a second experiment, these results were replicated by using inverted faces as a control condition, instead of misaligned faces (Schiltz & Rossion, 2006). More recently, they were replicated in an event-related (top) face identity paradigm, this time with concomitant behavioural measures of the composite face effect (Schiltz, Dricot, Goebel, & Rossion, 2010). Overall, the use of the composite face illusion in fMRI indicates that faces are represented holistically in face-sensitive areas of the visual cortex, in particular in the right hemisphere, providing a neural basis for the behavioural effects that have been described in many studies.

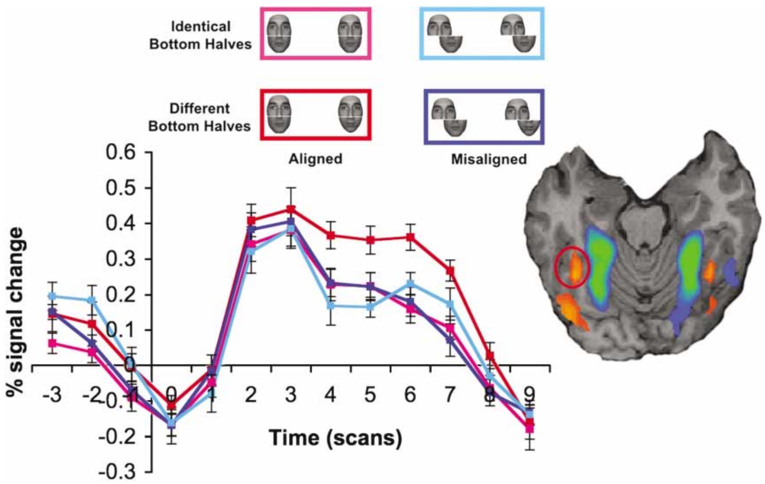

Figure 14. The composite face paradigm as used in the fMRI study of Schiltz and Rossion (2006). In the right fusiform face area (FFA), there was release from adaptation when the top halves were perceived as being of a different face identity (aligned with different bottom halves), a large effect compared to the 3 control conditions. To view this figure in colour, please see the online issue of the Journal.

Human electrophysiology. The composite face illusion has also been studied with event-related potentials (ERPs), not so much to inform about its localization, but in order to provide information about the time course of holistic face perception. Indeed, in the definition of holistic face processing (see earlier), one finds the notion that the different local parts of a face are integrated *simultaneously* into a global representation. Thus, according to this view, the *initial* representation of the individual face in the system should be holistic. ERP is an interesting technique for this endeavour because of its high-temporal resolution (Luck, 2005). In particular, we targeted the N170 ERP component, a large electrophysiological response peaking at about 170 ms following stimulus onset over occipitotemporal sites (Bentin, Allison, Puce, Perez, & McCarthy, 1996; see, for reviews, Rossion & Jacques, 2008, 2011). This visual component is interesting because it is associated with the initial activation of a face representation, and because it is the first response that is sensitive (i.e., reduced in amplitude) to the repetition of the same individual face, providing that the face is presented at upright orientation (Jacques, d'Arripe, & Rossion, 2007).

Similarly to the fMRI studies, top halves of faces with different aligned bottom halves produced larger N170 amplitudes than the same top halves of faces with the same bottom halves, as early as 160 ms poststimulus onset (Jacques & Rossion, 2009) (see Figure 15). Again, this early effect—which was abolished when misaligned faces were presented—concerned essentially the right hemisphere. Therefore, this study, replicated later (Kuefner, Jacques, Prieto, & Rossion, 2010), identified the functional locus of the composite face effect at the earliest face perception stage, suggesting that facial parts are not independently processed as face-like entities before being integrated into a holistic representation.

Convergent validity. While I've mostly illustrated the composite face studies performed by my colleagues and myself, many other researchers have also used this paradigm. Considering only the studies that focused on face identity, some studies have aimed at understanding the nature of the composite face effect by manipulating the stimuli (e.g., top face halves aligned with moving bottom halves: Khurana, Carter, Watanabe, Nijhawan, 2006; composite effects for profile faces: McKone, 2008; two halves of a composite face separated in, or slanted through, stereodepth: Taubert & Alais, 2009), and one study has shown that ingroup members are associated with a larger composite effect than outgroup members (Hugenberg & Corneille, 2009). Other studies report the abnormality of the effect in populations of human observers with a lack of early visual experience from one eye (Kelly, Gallie, & Steeves, 2012), as well as showing that it is unaffected in populations deprived of nonvisual inputs (deaf people; de Heering, Aljuhanay, Rossion, & Pascalis, 2012). Finally, the composite face effect has also been used to show holistic processing of conspecifics in nonhuman primates (spider monkeys: Taubert, 2010; Taubert & Parr, 2009; adaptation paradigm with eye movement recordings in rhesus monkeys: Dahl, Logothetis, & Hoffman, 2007; two alternative forced choice in rhesus monkeys and chimpanzees: Taubert, Qureshi, & Parr, 2012).

Although face processing is a field that is replete with debates and disagreements, it is important to note that the observations reviewed so far in this paper, and their conclusions, have been generally well supported by studies that used other paradigms to measure holistic face perception. To give an overview, for "other-race" face studies see, for example, Michel, Caldara, and Rossion (2006) and Tanaka et al. (2004); for low spatial-frequency dominance, see Goffaux (2009); for studies in children see, for example, Carey and Diamond (1994), Pellicano and Rhodes (2003), Tanaka, Kay, Grinnell, Stansfield, and Szechter (1998), and the review of Crookes and McKone (2009); for studies with cases of acquired prosopagnosia see, for example, Levine and Calvanio (1989), Sergent and Villemure (1989), and Van Belle, de Graef, Verfaillie, Rossion, and Lefèvre (2010b); for neuroimaging studies see, for example, Andrews, Davies-Thompson, Kingstone, and Young (2010), and Harris and Aguirre (2008); for studies reporting deviations from linearity at orientations around 90° when measuring holistic

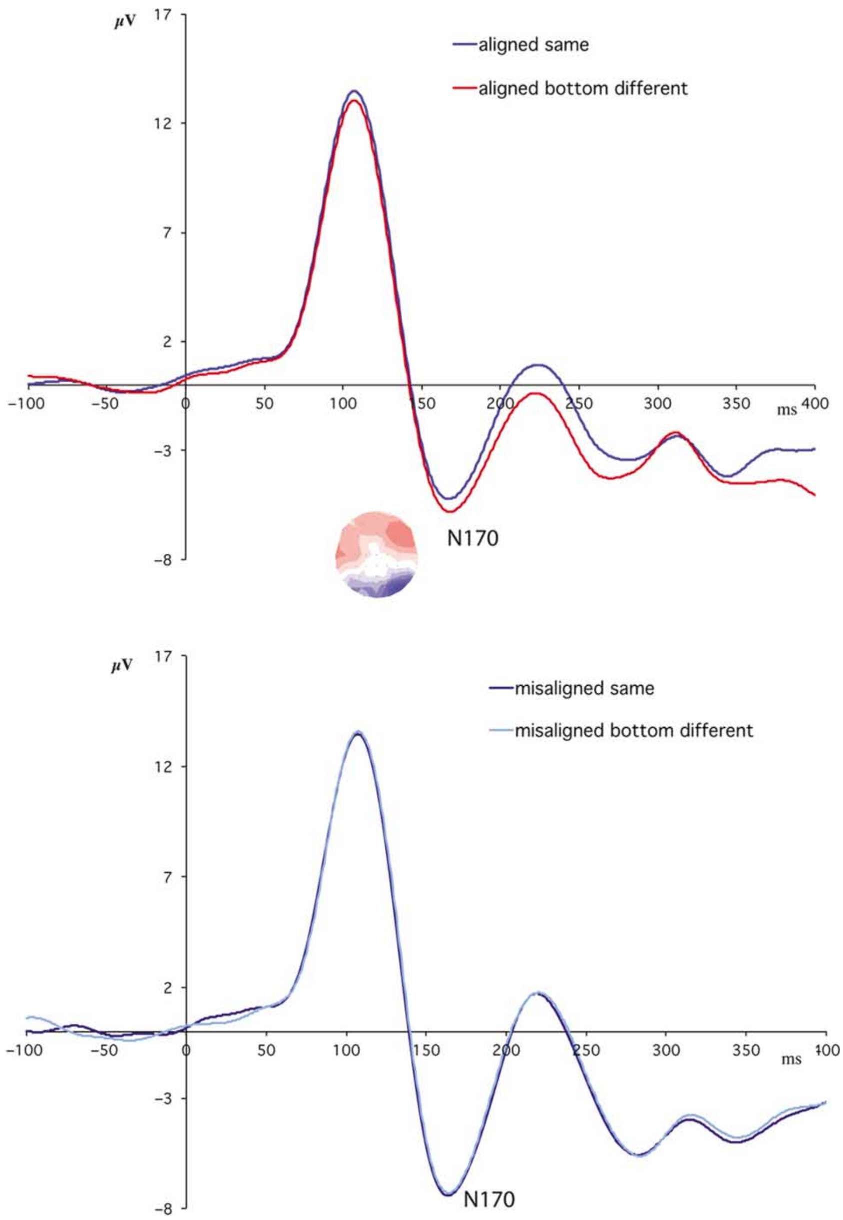

Figure 15. As early as the face-sensitive N170 to a test face preceded by a study face, there is a larger amplitude when identical top halves of the study and test face have different bottom halves than when their bottom halves are identical. This result is observed on right occipitotemporal sites (scalp topography of differential amplitude displayed, next to the N170), and only when faces are spatially misaligned (Jacques & Rossion, 2009; see also Kuefner, Jacques, Prieto, Rossion, et al., 2010). To view this figure in colour, please see the online issue of the Journal.

face perception see Lewis (2001), Murray, Yong, and Rhodes (2000), Sjoberg and Windes (1992), and Stürzel and Spillmann (2002); as well as other experiments aimed at testing the effect of orientation on holistic face perception using tasks such as the categorical perception of faces in noise (McKone, Martini, & Nakayama, 2001), the perception of a "Mooney" face stimulus (McKone, 2004), or the matching of "Thatcherized" faces (Edmonds & Lewis, 2007).

In summary, even though there remain some disagreements, often in the interpretation rather than in the data themselves (e.g., see Footnote 2 and the end of Part 2), the composite face paradigm appears to provide a robust and simple way of assessing perceptual integration between the parts (defined here as one half and the other half) of an individual face.

PART 2: THE MEASURE OF AN ILLUSION

The composite face paradigm

The second goal of this review is to explain the rationale behind the composite face paradigm, how to use it under different circumstances, discuss what can and cannot be inferred from it, and how to improve it. The reader may find it strange that I explain the rationale for this paradigm *after* reviewing the findings made with its use. However, this is also how the story unfolded: Following the studies of Young et al. (1987) with a recognition task and Hole's (1994) introduction of the matching task variant, experimenters used the matching paradigm extensively, without providing much theoretical and methodological justification for the conditions and parameters used in their studies. These studies have in common the use of composite faces, and the comparison of aligned and misaligned conditions. Nevertheless, they can differ greatly in terms of methodological parameters. Even in different studies from the same laboratory, the paradigm has been modified substantially, mainly in order to fit the technique used (behavioural, ERP, fMRI, . . .) and the population tested (e.g., children, patients with prosopagnosia) (e.g., compare the paradigm used in Michel, Rossion, et al., 2006 to Schiltz & Rossion, 2006, or to Kuefner, Jacques, et al., 2010, or to Jiang et al., 2011). Over the years, the composite face paradigm has even been modified in different behavioural studies performed in my own laboratory, simply because we tried to improve it—that is, make it tightly controlled and at the same time as sensitive as possible—progressively. For these reasons, it is important to have a good understanding of the paradigm and what it is supposed to measure in order to be able to account for putative discrepant findings, and optimize the paradigm for future studies. This second part should also provide the reader with all the information necessary to dispel the myth that the composite face paradigm is "partial" or "flawed", an issue that will be addressed in the third part of this paper.

The basic composite face paradigm. Despite these variations, the composite matching paradigms that have been used in the 60 or so studies reviewed above have one thing in common: In line with the original demonstration of Hole (1994, Exp. 1), they all consider that the important condition to use is one in which participants are asked to make a judgement of identity on two physically identical top halves that are aligned with two physically different bottom halves (Figure 16).

This is the critical condition of the paradigm. It is usually compared to a control condition in which the two face halves are spatially misaligned (Figure 17), so that the two top halves are now correctly perceived as being identical.

Engaged in a delayed matching task on these two conditions (Figure 4), observers perform less well in the aligned as compared to the misaligned condition (Figure 18).

The difference, in terms of accuracy rates and correct RTs, between matching the two identical top halves when their respective different bottom halves are aligned as compared to misaligned can then be taken as an index of perceptual integration of an irrelevant "part" (= bottom half of a face) with a target "part" (= top half of the face). The index of holistic face perception for faces is then computed as a simple difference:

"Same" trials: Performance (misaligned) – Performance (aligned)

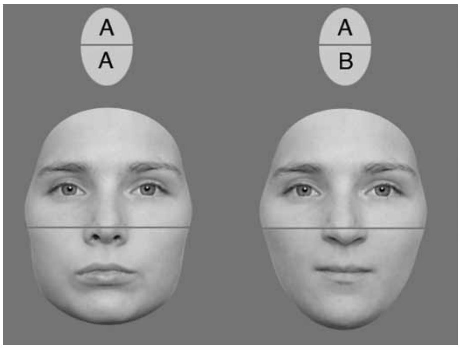

Figure 16. The two basic condition of the standard composite face paradigm. The two top halves (above the thin gap) are physically identical yet they are perceived as different because they are aligned with different bottom halves.

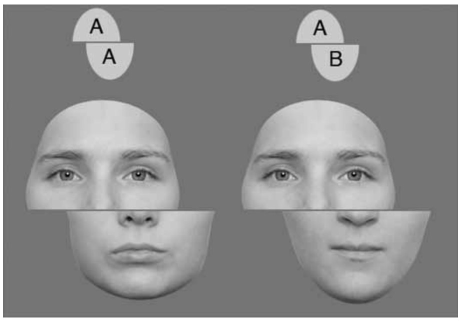

Figure 17. The same stimuli as shown in Figure 16, but with a slight spatial misalignment between the top and bottom halves. In this condition, despite the presence of different bottom halves, the observer has no difficulty determining that the two top halves are identical.

Note that an index of holistic face perception can also be computed as:

"Same" trials: [Performance (misaligned) – Performance (aligned)] / [Performance (misaligned) + Performance (aligned)]

Or as:

"Same" trials: [Performance (misaligned) – Performance (aligned)] / [Performance (misaligned)]

In all cases, the composite face effect is primarily an index of the consequences of *spatial (mis)alignment*.

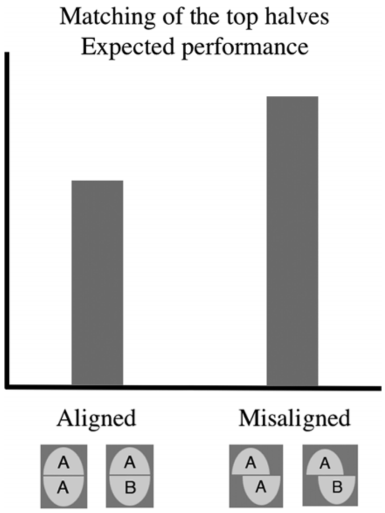

Figure 18. Expected performance on a composite face paradigm performed with the two conditions described earlier. Performance is measured in accuracy rates (higher for misaligned than aligned, as displayed here) but also in correct RTs (higher for aligned than misaligned).

Why misalignment?

Back to our Gestalist roots: (Lateral) spatial misalignment is theoretically relevant. Spatially misaligning, laterally, the top and bottom halves of a face breaks a powerful law of Gestalt perception: The law of *continuity*, or good continuation, which states that oriented units/points tend to be integrated into perceptual wholes if they are connected or aligned with each other in straight or smoothly curving lines (Pomerantz & Kubovy, 1986; Wertheimer, 1925/1967). Spatially misaligning the two halves of a face breaks this continuity of the contour of the face by introducing an edge, or a nonaccidental property (NAP; Biederman, 1987; Lowe, 1985). Thus, spatial misalignment corresponds to a *physical separation* of the whole face into parts. It is a small manipulation of the stimulus, but one that goes directly against perceptual integration of parts or elements (e.g., Altmann, Bülthoff, & Kourtzi, 2003). Moreover, spatial misalignment in the composite face paradigm does not only create a stimulus that cannot fit any internal holistic representation: Contrary to a vertical separation (a "gap", see later), segmenting the two parts by *laterally* moving the bottom part prevents the visual system to complete the contour of the face (the so-called Gestaltist law of *closure;* Pomerantz & Kubovy, 1986; Wagemans, Elder, et al., 2012; Wertheimer, 1925/1967). In short, there are good reasons why spatial misalignment is a theoretically relevant control manipulation, perhaps the best that one could come up with (credit to Young et al., 1987).

Breaking apart or increasing metric distances?. When comparing aligned to misaligned faces, one can safely attribute the difference in performance to a single manipulation: Spatial (mis)alignment of the parts. However, spatially offsetting the bottom half of a face from its target top half has at least two consequences. First, it *breaks the whole stimulus configuration*, changing dramatically the shape of the whole stimulus. Second, this manipulation also increases the *metric distance* between diagnostic features of the top half (e.g., the eyes) and of the bottom half (e.g., the mouth). Of course, if one uses relatively small stimuli, this increase of relative distances between facial features can be minimized, as we have attempted to do in most of our studies (see Figure 19). However, it remains the case that facial parts such as the mouth are located further away from fixation in misaligned than in aligned faces, and one could argue that this is the very reason why spatial misalignment disrupts the visual illusion and the composite face effect.

Behavioural studies reported so far in the literature cannot tell which of these two effects (introduction of an edge *or* increase of metric distance) is responsible for the fact that the bottom half is no longer integrated with the top half in misaligned faces. To address this issue, we recently designed an experiment in which the bottom half of the composite face was spatially shifted from the top half in parametrically increasing steps of 16% face width (Figure 20). We reasoned that if the loss of the composite effect for misaligned

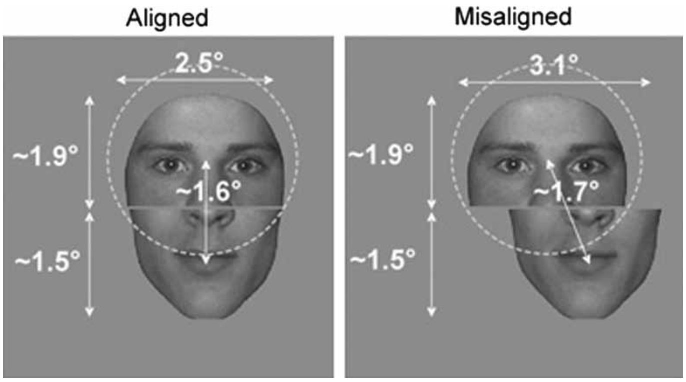

Figure 19. Examples of stimuli from a composite ERP study (Jacques & Rossion, 2009), showing that the increase of distance between the centre of the mouth and the eyes for misaligned as compared to aligned faces can be minimal. Yet, the distance between the eyes and mouth/chin increases for misaligned trials.

faces is due to an increase of distance between facial parts, then this composite effect should still be substantial for a minimal amount of spatial misalignment, and should decrease linearly with an increase in spatial misalignment (i.e., distance) between the top and bottom halves. Alternatively, if the effect is primarily due to the breaking of the whole face configuration, most if not all of the effect should disappear immediately with only a minimal amount of spatial misalignment (a change in nonaccidental properties), with no further (linear) increase associated with increasing degrees of spatial misalignment (a change in metric properties). The results of that study (see Figure 20; from Laguesse & Rossion, 2011) were crystal clear. Overall, there was a strong effect of misalignment, that is, a composite face effect. Critically, the effect of misalignment was fully accounted for by the difference between the aligned condition and all the other conditions: When the aligned condition was removed from the analysis, there was no longer any difference between the other conditions. This observation was made even when an additional minimal amount of spatial misalignment (8% face width) was included in the analysis. Thus, spatially misaligning the faces by 8% or 100% of face width did not make any difference (Figure 20) for the magnitude of the effect.

This observation reinforces the importance of the misaligned condition as a control in this paradigm because, if possible, one should always have a control condition that differs as little as possible from the condition of interest. More generally, this observation helps in clarifying the nature of the composite face effect/illusion: For upright faces, it is critically due to the spatial continuity[3] between the two face halves, so that they form a whole configuration.

[3] Except for the small gap in between the two halves, which does not prevent the composite face effect, an issue that will be discussed later.

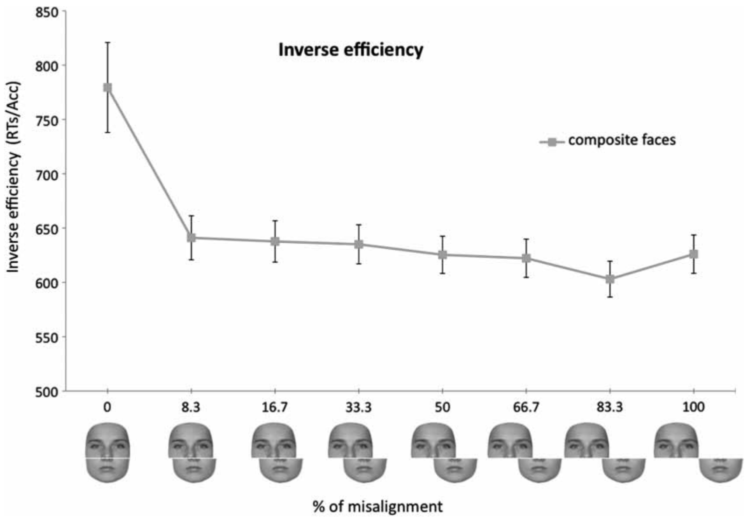

Figure 20. The magnitude of the composite face effect is independent of the length of misalignment between the two halves. A very small spatial misalignment of the two halves (8% of face width), introducing an edge, improves participants' performance (decrease of inverse efficiency) as much as if the two parts are completely misaligned (from Laguesse & Rossion, 2011).

Inversion (yet again). As mentioned previously, rather than using misaligned faces as a control for aligned faces, one can use the exact same aligned faces presented upside-down, as in Hole's studies for instance (1994; Hole et al., 1999). Predictions for such a condition are a little bit more difficult to make though, because human observers do not process inverted faces very well (Yin, 1969; see, for a review, Rossion, 2008). Hence, it may be that when they are upside-down, the top parts are not perceived as being identical so easily. However, paradoxically, in the context of the composite face effect, observers appear to perform better with inverted than upright faces (Hole, 1994; Hole et al., 1999; Young et al., 1987) (see Figure 21).

Controlling for general effects of alignment. In order to control for general effects of misalignment, one could also add conditions in which everything remains the same: The two identical top halves are aligned or misaligned with identical bottom halves (e.g., Busigny et al., 2010; Jacques & Rossion, 2009; Jiang et al., 2011; Ramon et al., 2010, Exp. 5; Schiltz et al., 2010; Schiltz & Rossion, 2006; see Figure 22). The index of holistic face perception can then be computed as:

"Same" trials: Same bottom half [Performance (misaligned) – Performance (aligned)] – Different bottom half [Performance (misaligned) – Performance (aligned)] (i.e., an interaction).

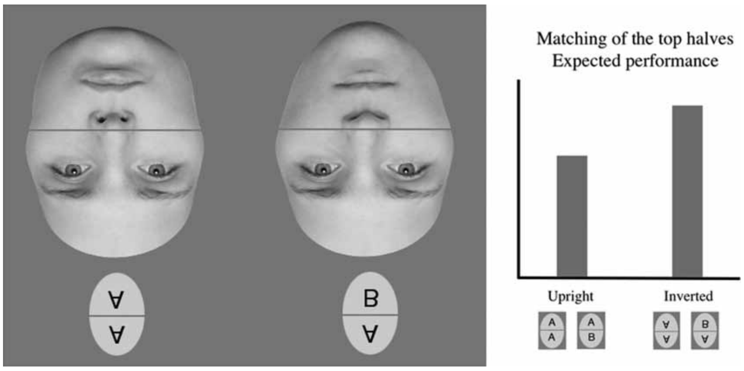

Figure 21. The same stimuli as shown in Figure 16, but presented upside-down (flipped vertically). In this condition, the observer usually has less difficulty telling that the two "top" halves are identical, despite the presence of different "bottom" halves, and the fact that the stimuli are presented upside-down.

Note that adding such conditions, in which everything remains the same (Figure 22, or left test faces on Figure 23), does not substantially change the magnitude of the composite face effect, as illustrated by data from our recent studies (Busigny et al., 2010; Jiang et al., 2011). The reason being, the performance for trials in which the bottom half does not change (Figure 22) is not much influenced by the spatial alignment of the two halves (Figure 23).

Therefore, in these behavioural studies, as shown in Figure 23, one can directly compare the aligned and misaligned conditions for which the bottom halves differ, as in the standard composite face paradigm. However, when using composite faces in fMRI (Schiltz et al., 2010; Schiltz & Rossion, 2006) or ERPs (Jacques & Rossion, 2009), it might be important to add trials in which everything remains the same in order to remove any general effect of spatial alignment on neural activity. For instance, simply spatially misaligning two halves of a face picture paradoxically increases the face-sensitive N170 component (see Letourneau & Mitchell, 2008; Jacques & Rossion, 2010; see Figure 15). One can control for that effect by including trials in which both the top and bottom halves remain the same between the two faces to compare. Moreover, these studies are performed in the context of a face adaptation paradigm, in which a stimulus with a change of property is typically compared to a fully repeated visual stimulus (baseline).

"Different" trials. Finally, since the behavioural task in the composite face paradigm is a "same/different" task, trials requiring a correct "different" response should also be included. However, "different" trials do not lead to any composite face illusion: Two physically different top

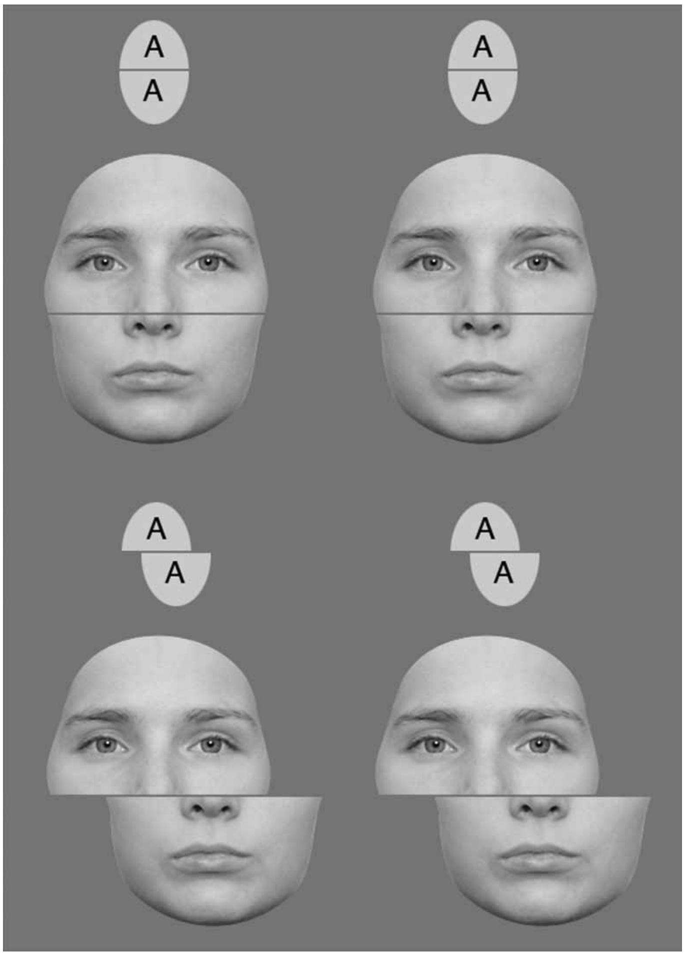

Figure 22. Here both the top and bottom halves are strictly identical, so that adding these 2 conditions provides an additional baseline, or control, in the paradigm.

halves of faces do *not* look more similar when they are aligned than when they are misaligned with identical bottom halves (Figure 24). Therefore, there should be no composite face effect when having to differentiate two different top parts of a face that have the same bottom halves.

To prove this claim let me just show some data obtained from 24 participants of a study (Gao, Flevaris, Robertson, & Bentin, 2011) that included different top halves associated with identical bottom halves (as in Figure 24). For "same" trials with different bottom halves, performance is better for misaligned (92.2%) than aligned (79.5%, black column) trials,

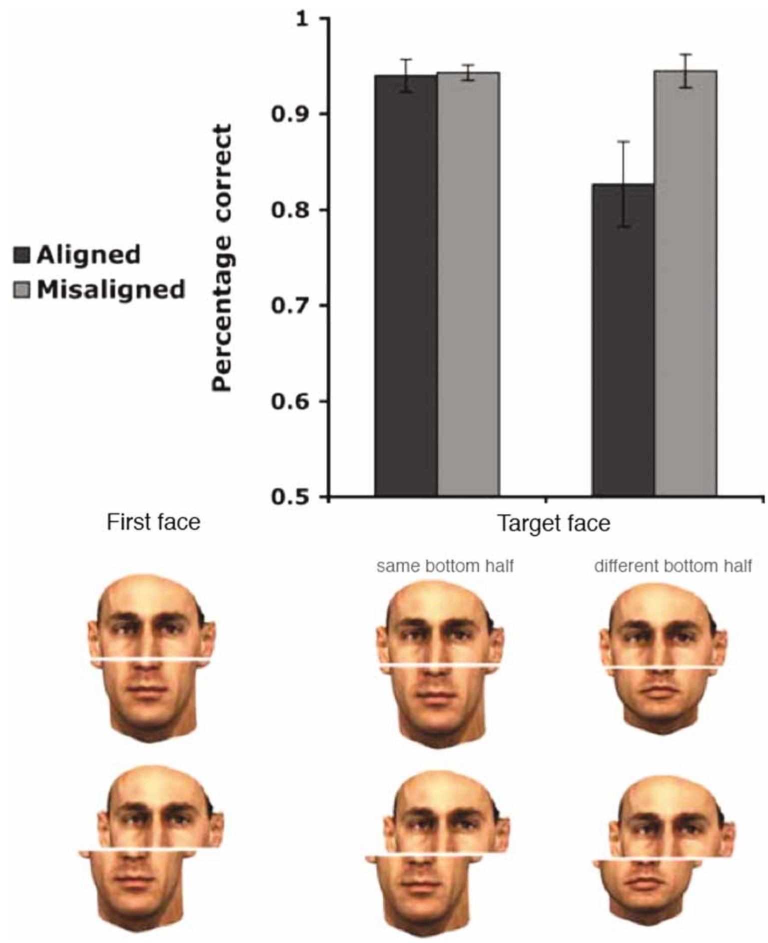

Figure 23. Data of a delayed matching task with composite faces (figure adapted from Figure 3 of Jiang et al., 2011), showing the decrease of performance for aligned as compared to misaligned faces, only when the bottom halves are different between the two faces. There is no effect of alignment when the bottom halves are identical. To view this figure in colour, please see the online issue of the Journal.

(12.7% difference), $t(23) = 38.41$, $p < .0001$ (Figure 25). This is how the composite face effect is typically measured, at least for accuracy rates. In contrast, for "different" trials with identical bottom halves there is no difference between aligned and misaligned conditions (92.36% vs. 93%), $t(23) = 0.38$, $p = .39$. These comparisons provide a direct demonstration that if the data is acquired correctly (i.e., if the participants used only the target half to make their judgement), "different" trials are not relevant in the computation of the composite face effect.

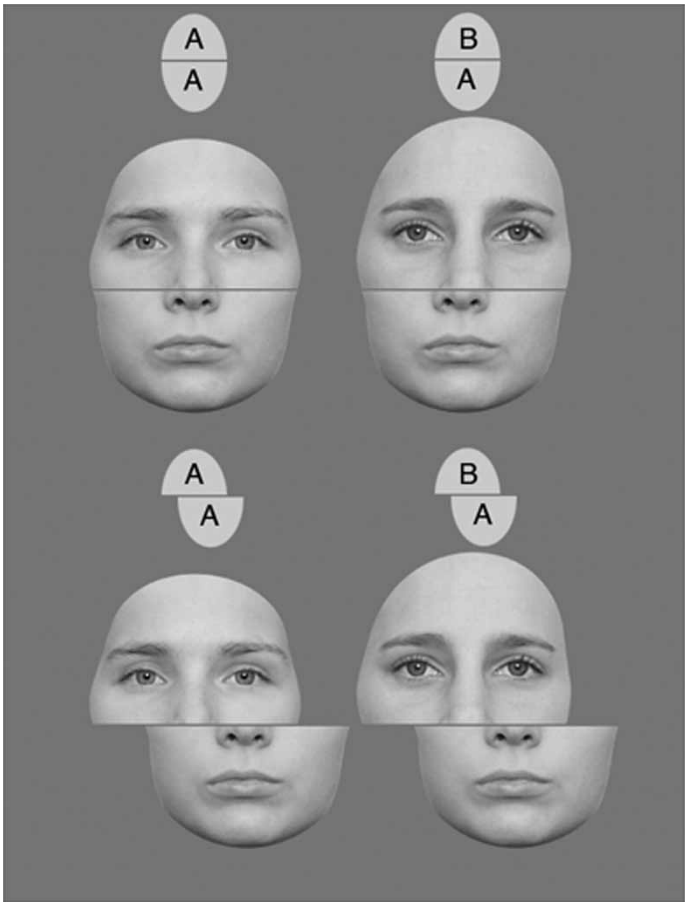

Figure 24. Here the top halves, selected from the faces used in Figures 16 and 17, are physically different, but their bottom halves are identical. One does not perceive these top halves as being more similar to each other when the two halves are aligned (top) than when they are spatially misaligned.

In fact, importantly, if an observer uses the *whole* face to do the task, he/she will always respond "different" even with identical bottom halves in aligned "different" trials (Figure 24). It is only if the observer uses the bottom half independently of how the whole face looks like (i.e., a part-based judgement), that he/she would respond "different" in such trials. For this reason, "different" trials with identical bottom halves should never be included in the paradigm. Rather, "different" trials should have both halves as being different (Figure 26). Ideally, data on these "different" trials should be reported in the paper—independently of the "same" trials—so that one can verify that there were no unexpected effects of misalignment (for

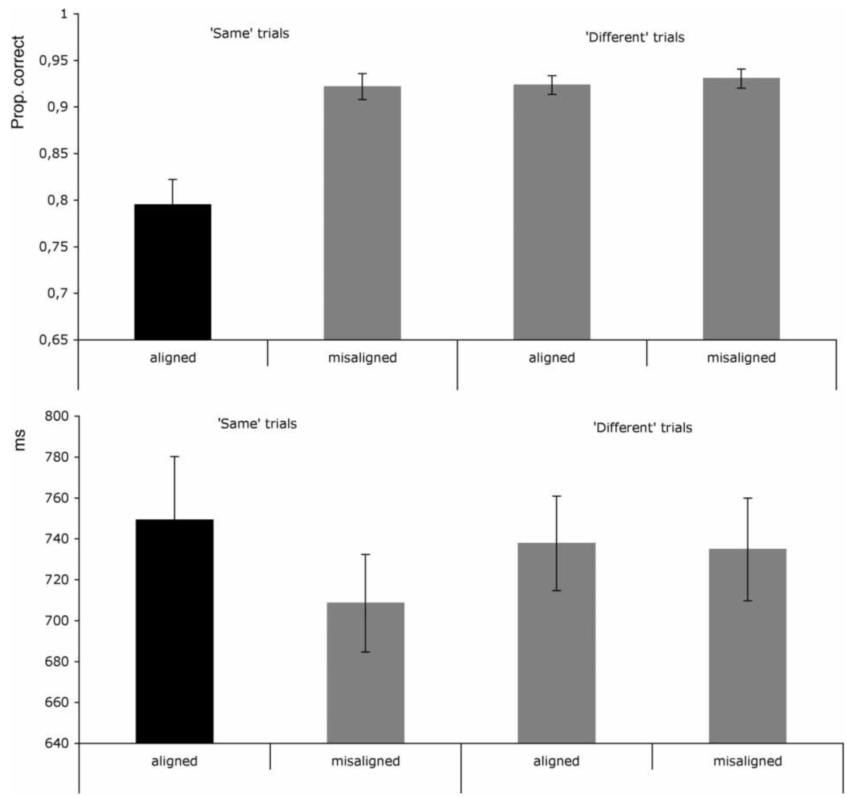

Figure 25. Data (24 participants) of the study of Gao et al. (2011) reanalysed by separating the trials usually included in the standard composite face paradigm ("same" trials: Identical target top halves with different bottom halves) from the "different" trials (different target top halves associated with identical bottom halves). (A) The composite face effect is illustrated by the reduction of accuracy rates for the "same" aligned trials (in black) relative to misaligned trials. On the right side of the figure, one can see that there is no such effect for "different" trials. These trials are not associated with any composite face illusion (Figure 24), and thus are largely irrelevant. (B) Correct response times (trials below 200 ms removed). Note again the increase of response times in the aligned condition for "same" trials, which is absent for "different" trials. There was a significant difference between aligned and misaligned "same" trials, as used in most studies to assess the composite face effect, $t(1, 23) = 12.2$, $p < .002$, but no such difference for "different" trials, $t(1, 23) = 0.05$, $p = .83$.

instance, a general trend to press "same" for misaligned and "different" for aligned trials). Alternatively, they could also be used together with "same" trials in the computation of the effect (see later), but only if both halves of the "different" trials differ.

Why "different" trials do not give rise to a composite face illusion/ effect?. Why is it that "same" but not "different" trials lead to a composite illusion/effect? Let me speculate about a number of potential reasons.

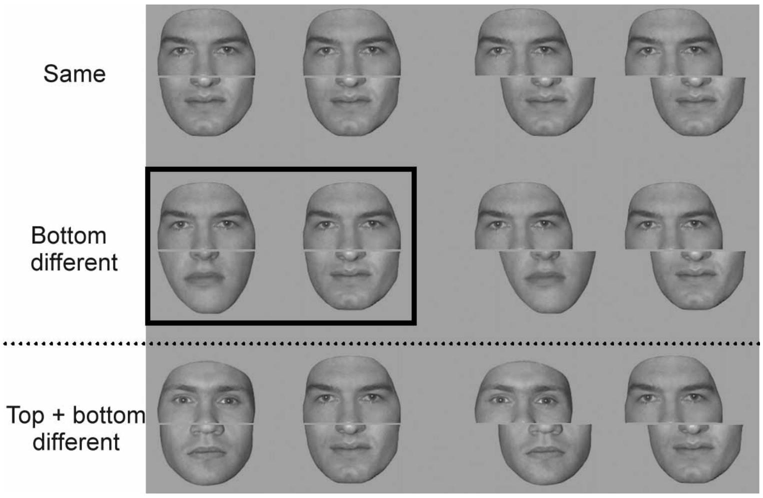

Figure 26. The composite face paradigm as used in several of our studies (stimuli and adapted figure from Jacques & Rossion, 2009; see also Busigny et al., 2010; Jiang et al., 2011; Ramon et al., 2010, Exp. 5; Schiltz et al., 2010). Because "different" top halves with identical bottom halves are not associated with a composite face illusion, the "different" trials (below the dotted line) in the paradigm usually differ by both the top and bottom halves. They are irrelevant for the computation of the composite face effect. The critical conditions (in the rectangle) are the ones used in the standard composite face paradigm in which the two top halves are identical but are perceived as different when they are aligned versus misaligned.

First, in general, it is easier to create a (wrongly perceived) difference between two stimuli that are identical than to create the same identity from two stimuli that are different. To illustrate this point, one can use a visual illusion simpler than the composite face illusion, such as a variation of the classical Müller-Lyer illusion (Müller-Lyer, 1889). It is easy to take two arrows of equal length ("same trials") and make them erroneously perceived as being of different lengths (Figure 27A). However, in order for two arrows of different lengths ("different trials") to be erroneously perceived as having the same length (Figure 27B), the stimuli have to be carefully adjusted. Therefore, when the arrow heads on both sides of the lines are different, one would expect an increase in errors and RTs that is larger for pairs of arrows that are identical ("same trials", judged as different) than for pairs of arrows that are physically different ("different trials", judged as identical).

Second, while the stimulus of the Müller-Lyer illusion is unidimensional, faces are *multidimensional*: Individual faces can vary on a very large number of cues that can be diagnostic of face identity. These cues concern the shape and the surface (colour, texture) characteristics of the face, which can vary at the local (i.e., specific parts) and global (i.e., overall shape, skin texture) levels. In an influential account, these diagnostic cues for individualizing

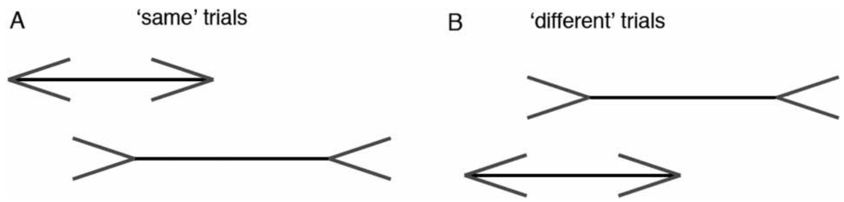

Figure 27. Two versions of a variation of the Müller-Lyer visual illusion. (A) The shafts of the arrows are of equal length, but the arrow with diverging heads is seen as longer than the one with converging heads (a "same" trial, giving rise to a "different" response, as in the composite face effect). (B) The shafts of the arrows are of different lengths, the arrow above being shorter than the arrow below. Yet, these two look roughly identical, or at least much closer in length than the two arrows on the left.

faces have been conceptualized as dimensions in a face-space (Valentine, 1991). Even without considering the human observer, these cues, or dimensions, are not independent from each other: Variations along one dimension can be systematically associated with variations in other dimensions, so that the number of dimensions or factors accounting for most of the variance between individual faces is not that large, and can be estimated by principal component analysis (PCA; e.g., O'Toole, 2011; Turk & Pentland, 1991).

If two faces differ across multiple dimensions, as in "different" trials, judging that they are different is straightforward. Especially if the dimensions on which the faces differ are fixated, and even more so if the faces differ at the level of the eyes region, the most diagnostic region of the face (e.g., Gosselin & Schyns, 2001; Haig, 1985; Sadr, Jarudi, & Sinha, 2003). In fact, judging a difference between two faces may potentially rely on the detection of a single local cue, such as the colour of a single eye (Figure 28).

Because of that, one does not need to go beyond a *single* local cue to make a "different" judgement on two top face halves. It does not matter if many other cues, outside of fixation, are identical between the faces to match: They will not make the faces look more similar to each other. In fact, even if the two faces are identical on all dimensions but one, they will readily look different (Figure 28). As Galton (1883, p. 3) put it in his insightful discussion of holistic face perception, "one small discordance overweighs a multitude of similarities...". In other words, considering a *single* cue is sufficient to make the "different" judgement, even a cue that can be relatively independent of other cues (i.e., the change of one eye colour as in Figure 28). Therefore, although normal observers certainly rely on holistic processing when discriminating unfamiliar faces, the integrity of holistic processing is not strictly necessary to perform such a discrimination task, especially if the diagnostic cue is fixated. Discriminating faces that differ in a fixated face part neither requires nor promotes holistic face perception. In contrast, almost by

Figure 28. "Watch that man". A single difference, the colour of the right eye, is extremely salient and can make the whole face look different. The author has obtained copyright permission from the photographer to use the image. To view this figure in colour, please see the online issue of the Journal.

definition, judging whether two half faces are exactly identical requires that *all* the cues of the fixated half face, as well as their *relationships* (e.g., relative distances between the eyes and other so-called "configural cues") are taken into consideration. Thus, even though a "same" judgement could be performed analytically, considering one part after the other, this kind of judgement inherently *promotes* the opportunity for holistic processing to be observed in the task. This is the second reason why "same" trials are relevant in this composite face paradigm, whereas "different" trials are not relevant.

To summarize this section, even when considering only the stimulus conditions (not the other methodological parameters), the classical composite face paradigm can have several variants. However, it is generally a simple, well-balanced paradigm with two important basic conditions that differ only by a single factor (spatial misalignment of parts). The observer makes more mistakes (i.e., increases the rate of "different" responses) and/or takes longer at matching two identical top halves of faces in an aligned condition as compared to the exact same condition in which the two face halves are (slightly) spatially misaligned. This misalignment effect obtained on "same" trials with different bottom halves is taken as a behavioural index of perceptual integration, or holistic face perception.

Should Signal Detection Theory be used to analyse the composite face paradigm?. Signal Detection Theory (SDT; Green & Swets, 1966) has been developed to deal with situations in which there is a signal to detect against background noise, as in a yes (signal)/no (noise) paradigm. An advantage of SDT is that it makes full use of a participant's responses by combining information from both the "yes" and the "no" trials to compute a measure of sensitivity (d') and of response bias/criterion rather than relying only on accuracy rates. SDT can be useful in a typical old/new recognition task, for instance, because it takes into account differences between performance when there is a signal (a face seen previously) and when there is no signal (a face never seen before). If there is a signal and the observer detects it ("yes" response), then it is recorded as a *hit*. If, despite the absence of signal, the observer provides a "yes" response, this is considered as a *false alarm* (FA). If there is only noise and the observer does not detect any signal, this is a *correct rejection*. If there is signal and the observer misses it ("no" response), it is a *miss*. In the context of an old/new face recognition task, two participants might present with the same performance in terms of overall accuracy rates (e.g., 70%), yet they may behave very differently: One participant may recognize all faces seen previously but also have a high rate of false recognition ("hyperfamiliarity"), whereas another participant may correctly reject all unknown faces but fail to recognize a fair number of faces seen previously ("low familiarity", which is typical of cases of acquired prosopagnosia). Although the d' measure will be identical in the two situations, the direction of criterion/bias will be opposite. Therefore, in such a task, SDT offers a richer assessment of an observer's behaviour than accuracy rates alone (Stanislaw & Todorov, 1999).

However, in the composite face paradigm, the only trials that matter are the "same" trials. No difference is expected between aligned and misaligned "different" trials because both face halves (must) differ in these trials. Rather than using accuracy rates on "same" trials only, one can nevertheless use SDT in the composite face paradigm to compute a d' and a *bias/criterion* by measuring the proportion of *hits* on "same" trials, and the proportion of FAs on "different" trials, separately for aligned and misaligned condition. This procedure allows assessing the discrimination performance in the task. However, there are three caveats associated with this procedure. All three reflect important aspects of the paradigm, which deserve to be discussed in independent sections.

Looking for a response bias. Because "same" trials, but not "different" trials, carry the composite face effect, there is an inherent *response bias* in the composite face paradigm: The number of "different" responses is artificially increased. For instance, if there are 50% "same" trials in the paradigm in total, one may observe 55–60% "different" responses in total. Thus, the standard composite face paradigm is a paradigm *that is intended to cause a response bias*: In the critical condition ("same" aligned trials), the top halves of faces are perceived as slightly different despite being identical. Hence, the proportion of "different" responses will be higher than expected in this condition. This response bias is exactly what experimenters aim for in this same/different composite face paradigm: People's perception is fooled and it leads them to increase artificially their proportion of "different" trials. However, critically, this response bias is expected only in the aligned condition, not in the misaligned condition.

A response bias of perceptual origin. Using SDT to analyse the composite face paradigm leads to two variables, the *d′* and the bias/criterion. Which of these variables should be used to assess the magnitude of the composite face effect? Some authors may consider the bias/criterion as being a "*response* bias" or a "*decision* bias", i.e., an effect of "cognitive/decisional" nature, whereas the *d′* measure would reflect "true discriminability" (i.e., an effect of perceptual origin). However, such an interpretation can be profoundly misleading because the bias/criterion could be as valid a measure as the *d′*, and have a perceptual basis. That is, SDT is agnostic about the origins of the bias/criterion, and cannot inform about the functional locus of an effect (i.e., perceptual, attentional, decisional/response, . . .). Indeed, biases of perceptual origin exist and are readily induced. For instance, prolonged exposure to a moving stimulus leads to perceived illusory motion of static stimuli, the motion aftereffect (Nishida & Johnston, 1999; Wright & Johnston, 1985; often referred to as the "waterfall illusion", S. Thompson, 1880). It is undoubtedly a *perceptual* phenomenon: Following adaptation to motion, motion is *seen* on a stationary pattern, and direction-selective neurons in a visual area (MT+ complex) lower their firing rate as a function of adapting to motion in their preferred direction (Van Wezel & Britten, 2002). Yet, the signature of the motion aftereffect in psychophysical data analysed with SDT is a shift in the psychometric function, indistinguishable from "response bias" (Nishida & Johnston, 1999; Van Wezel & Britten, 2002). Hence, dismissing a difference in response bias between aligned and misaligned conditions, or interpreting it as reflecting an effect of decisional nature, may just be missing the whole point. For this reason, one should be careful when using SDT to analyse data in the composite face paradigm.

Composite face effects can arise without a behavioural same/different response (bias). The composite face effect has been shown in other paradigms than the widely employed same/different matching task. For instance, Young and colleagues (1987) reported a *naming* disadvantage (in RTs) for aligned face halves as compared to misaligned face halves (or inverted faces in their Exp. 2) and these authors concluded that holistic perception of face identity is a powerful perceptual phenomenon. Carey and Diamond (1994) also found significant composite effects in a naming task for familiar or experimentally familiarized faces, an effect recently replicated for personally familiar faces (Ramon & Rossion, 2012). Other studies reported the composite face effect in two-alternative forced choice tasks (Macchi Cassia et al., 2009; Taubert et al., 2012; Turati et al., 2010). Laurence and Hole (2012) showed a composite face effect in the context of a face identity adaptation paradigm, in which participants" prolonged viewing of the composite was decoupled from their response to it, measured much later in the course of the experiment. Finally, the "neural" and "electrophysiological" composite face effects have been observed in paradigms that did not require same/different decision tasks or even the processing of facial identity (Schiltz & Rossion, 2006; see also Kuefner, Jacques, et al., 2010, for a composite face effect in ERPs without any behavioural response). In all these examples, there is no response bias. They serve to illustrate that the composite face effect does not necessarily depend on a response bias, but that the typical response bias is a consequence of the same/different composite face paradigm in which the effect is expected only for "same" trials.

Proportion of "same" and "different" trials. Another reason why SDT may not be recommended when analysing data of the composite face paradigm is that the proportion of "same" and "different" trials is not always equal in the paradigm, for good reasons. For instance, we tend to use a reduced proportion of "different" trials in our studies (e.g., 42% in Michel, Rossion, et al., 2006; 33% in Michel et al., 2007; Busigny et al., 2010; de Heering et al., 2007; Ramon et al., 2010). In the study of Rossion and Boremanse (2008), an experiment that included seven orientations for each condition (see Figure 10), there were only 29% "different" trials in the paradigm. For another reason, in the fMRI experiments of Schiltz and Rossion (2006), the top halves of faces were the same in 100% of the trials (see Figure 14). In that study, participants only had to detect a colour change on these top halves so that there was no need to include conditions with different top halves.

Besides reducing the duration of the experiment, using a smaller proportion of "different" trials has two other advantages. First, since participants are unaware of the different proportions, it can only increase their tendency to respond "different" for "same" trials, leading to a higher proportion of mistakes and the chance to observe more clearly the composite face effect (i.e., larger increase in "different" responses for "same" trials for aligned rather than for misaligned faces). Second, if one includes many "different" trials in the study, participants might consider that in comparison to these real "different" trials, the illusory different top halves of faces do not look that different after all. Having a large proportion of "different" trials in the composite face paradigm, especially if the individual faces are very different from each other, might therefore lead to a reduction of the effect. For this reason, but also to avoid pixel-by-pixel comparisons, one can systematically change the size and/or position of the faces to match (e.g., Schiltz & Rossion, 2006), so that even in the "same" trials, the top halves to match are physically different.

Finally, note that in some nonbehavioural studies, it could be interesting to compare "same" and "different" trials, not to measure the composite face effect, but for other reasons. For instance, in the fMRI study of Schiltz et al. (2010), the aim was to go beyond the finding of a release from fMR-adaptation due to the composite illusion ("same" top halves perceived as different), and assess the magnitude of this effect compared to when the top half faces were truly physically different. Such a comparison is correct only if the signal results from an average of the same amount of trials for the two conditions. The same reasoning applies to the ERP study of Jacques and Rossion (2009; see Figure 15) testing the N170 face adaptation effect with composite faces. In both studies, the proportion of "different" trials was only one third, so that with six conditions (Figure 24) the number of "same" trials with different bottom halves was the same as the number of "different" trials.

In summary, rather than obeying a fixed rule, the proportion of "same" and "different" trials in a composite face paradigm could be, and should be, flexibly tailored to the needs of the experiment, with several factors worth being considered. What matters is that the proportion of "same" trials is identical for the aligned and misaligned conditions. In future behavioural studies, it would be useful to parametrically manipulate the proportion of "different" trials, to test the effect of this variable on the size of the composite face effect.

The importance of response times. Starting with the study of Young et al. (1987), numerous studies have reported the composite effect in correct RTs (e.g., Hole, 1994, Exp. 2; Carey & Diamond, 1994; de Heering et al., 2007; de Heering & Rossion, 2008; Rossion & Boremanse, 2008; Wang et al., 2012; see also Figure 31). In fact, RTs are sometimes the main or even the only variable that gives rise to significant composite effects (e.g., Hole, 1994, Exp. 2; Wang et al., 2012). This is a third reason why SDT is limited when analysing the composite face paradigm: It is unclear how this analysis should be applied to RTs, let alone combined with accuracy rates in an efficiency measure for instance (Townsend & Ashby, 1983).

What are the factors that may determine whether the effect is observed in RTs or error rates? In principle, if presentation duration is long, the effect is more likely to be found in RTs, whereas very short presentation times will tend to provide significant effects in accuracy. Most studies use relatively long stimulus durations (several hundreds of milliseconds) and disclose effects in accuracy and RTs, or RTs alone. In our studies, we have generally observed that participants' strategies differ in the composite face paradigm, with some being more conservative than others in making their judgement. Despite perceiving the top halves of the faces as being slightly different when they are aligned with distinct bottom halves, these participants attempt to control their response, and make sure that they can correctly respond "same" on these trials. Therefore, these participants will perform almost at ceiling in the accuracy portion of the task, but they will usually take more time for these aligned trials than for misaligned trials, which gives rise to an effect only for correct RTs. In any case, because RTs usually give rise to a significant composite effect, a full consideration of this variable in the composite face paradigm is critical. In fact, because the effect can be observed on different dependent variables in different participants, it is useful to assess speed–accuracy tradeoffs and get a complete measure of the magnitude in the composite face effect by combining accuracy rates and RTs in an inverse efficiency measure (e.g., Figure 20).

Top and bottom. In most studies, and in all the examples discussed in this review, the composite face paradigm uses the top half of the face as the target. The bottom half is used as the irrelevant half. It is worth noting that the composite face effect can be observed for the bottom half of the face as well. Young et al. (1987) observed a delay in identifying bottom halves of famous faces when they were aligned with top halves from other faces than when they were spatially misaligned. However, the effect was not as large as when using the top face halves, despite the fact that the "bottom half" in Young et al.'s stimuli, being defined as everything just below the eyes, encompassed a much larger part of the face (i.e., the whole nose) than the top half (see Figure 1 in that study). Considering only the studies that used the composite paradigm, there are very few delayed matching studies that have used the bottom halves as targets (e.g., Nishimura, Rutherford, & Maurer, 2008). For instance, we used the bottom halves as targets in an experiment with a case of prosopagnosia (Ramon et al., 2010, Exp. 4) because this patient had a tendency to focus on the lower part of the face. A composite face effect was found in normal observers, but it was of lower magnitude than when the top halves were used as targets in another experiment (Exp. 3) of that study. Besides the fact that the composite effect on the bottom half is generally weaker than on the top half, there are a few other reasons why the top half is, and must be, favoured as a target in such studies.

First, as already noted, people are more accurate at identifying faces from features located in the top half than the bottom half of the face (e.g., Davies et al., 1977; Gosselin & Schyns, 2001; Haig, 1985, 1986; Sheperd et al., 1981). Second, and most importantly, the composite visual illusion is much less striking, or not present at all, for bottom halves than for top face halves (Figure 29). Hence, if one aims at capturing the perceptual phenomenon in a behavioural (or neural) measure, it is worth using the illusion at its best, that is, when identical top halves are associated with different bottom halves, not the contrary.

Why is there such a top/bottom asymmetry of the composite effect/illusion? One may be inclined to believe that it is because information in the

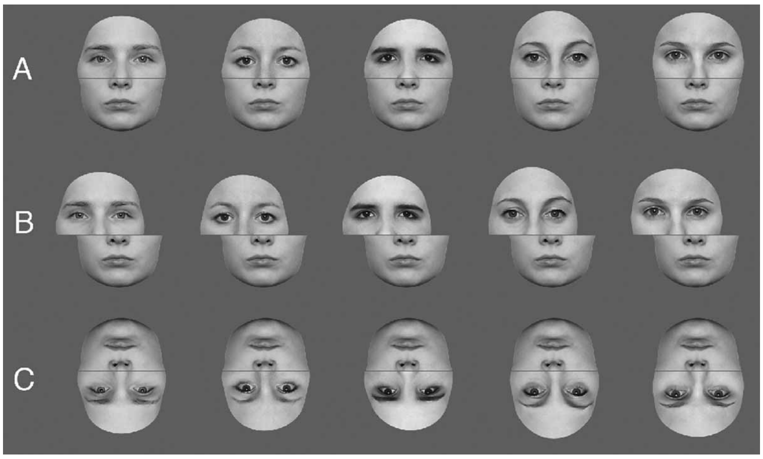

Figure 29. These are the same original stimuli as in Figure 1, although here the bottom halves are all identical, whereas the top halves are different. If you concentrate on the bottom halves in A, do you have the impression that they are different from one another? It is certainly not compelling. However, our attention is attracted by the top halves, which are physically different. Compared to Figure 1, this display suggests that the composite face illusion works essentially in one direction: When judging top halves but not when judging bottom halves. Here, B and C show the control conditions, in which the top and bottom halves are physically misaligned or the face is inverted, respectively.

top half of the face, the eyes and eyebrows in particular (when the hairline is not present), is usually dominant for individual face recognition. However, if anything, this factor should *reduce* the composite effect because if information on the top half alone is sufficient for matching identical top face halves, it is even more spectacular that the less diagnostic facial information in the lower part of the face influences performance in the composite face paradigm. Another reason might be that, under natural (unforced) circumstances, the location of the optimal fixation for face recognition appears to be central, slightly below the line of the eyes (Bindemann, Scheepers, & Burton, 2009; Hsiao & Cottrell, 2008). This fixation has been associated with the "centre of mass" of the face, and holistic face perception (Orban de Xivry, Ramon, Lefèvre, & Rossion, 2008; see Figure 30). Such a fixation point would allow that all facial features are well perceived, with a biased location on the top half of the face reflecting the larger number of elements, and increased saliency (Itti & Koch, 2000), of the top as compared to the bottom face half. Consequently, in a composite task, forcing an observer to fixate on the bottom half of the face, an unnatural fixation, may reduce holistic face perception. Moreover, having to fixate the top halves of faces is something natural for human observers, and there is a better chance that they respect that instruction. Although this is the case when participants are instructed to fixate the top half in the composite face paradigm (de Heering et al, 2008), it remains unclear if they are able to go against their natural fixation pattern and keep fixation on the bottom face half if they are required to.

A third factor that may account for the asymmetry between the top and bottom halves in the magnitude of the composite face effect is that the top half of the face contains more elements (two eyes, eyebrows, . . .) than the bottom half, which contains mainly the mouth as a salient feature. Therefore, the diagnosticity of the top half might be more *dependent* on the integrity of holistic perception, that is, the ability to see the many elements of a face as an integrated representation. This argument was developed when attempting to account for prosopagnosic patients' over-reliance on the mouth rather than the eyes region (Caldara et al., 2005; Orban de Xivry et al., 2008; Rossion et al., 2009; see Figure 30).

In any case, irrespective of the reasons behind the dominance of the composite face paradigm for the top halves of the face, it makes sense to use the composite face paradigm only with the top half in most cases. If one uses

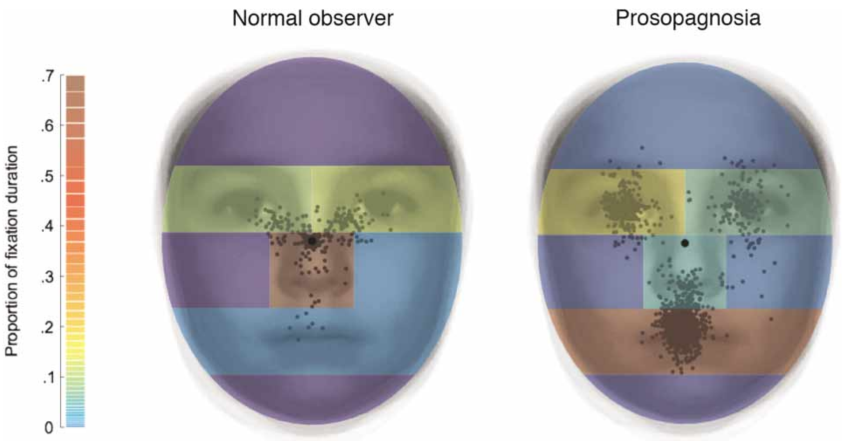

Figure 30. This figure (adapted from Orban de Xivry et al., 2008) shows the distribution of gaze fixations during a personally familiar face recognition task for a normal observer and for a well-known case of acquired brain prosopagnosia (PS; Rossion et al., 2003). During face recognition, a normal observer tends to fixate on the centre of the face, slightly below the eyes, rather than on any of the specific parts of the face (see also Hsiao & Cottrell, 2008). This fixation location is biased towards the superior half of the face, probably because of the larger number of diagnostic elements on—and higher saliency of—the top half of the face. This fixation location is thought to reflect the centre of gravity, or centre of mass for face recognition, and may be optimal for holistic face perception (Orban de Xivry et al., 2008). This view is supported by contrasting the fixation locations of a patient with acquired prosopagnosia (PS) who is impaired at holistic face perception: The patient rather fixates exactly on each part of the face, with a large proportion of fixations on the mouth (here, 60%), but also exactly on each eyeball (see also Van Belle et al., 2011). To view this figure in colour, please see the online issue of the Journal.

both the top and bottom halves in the paradigm, the asymmetry of processing and natural fixation between the two advises strongly against mixing up the top and bottom halves trials in the analysis.

Mind the gap. In some studies, a slight gap between the two face halves is included, as in all the illustrations of the present paper, and in fact in all of the studies that I coauthored. In the vast majority of studies though, such a gap is not included, most notably in the original study of Young et al. (1987). At first glance, adding a physical separation with a gap could only have a negative impact on the integration of the two facial parts. It even seems strange that one would want to separate artificially the two halves of a face when aiming to measure an integrated perceptual representation. However, although in an ideal situation, there should be no gap between face halves in order to maximize perceptual integration, I would like to argue that including a gap between the top and bottom half of the face in this paradigm is important. Indeed, if there is no gap in the aligned condition, how do participants of a study know what the experimenter exactly means by matching "the top half" of a face? This is problematic because, without a gap, the participant may consider that the top half includes the whole nose and he/she may attempt to match two top halves that contain some information that *is* physically different (e.g., the lower part of the nose). The paradigm may then lead to "different" responses for same aligned trials, even in the absence of perceptual integration. Moreover, researchers compare the aligned face condition to a misaligned face condition, in which the spatial misalignment provides exact information about the borders of the top and bottom halves. Therefore, without using a gap, the composite face effect could be artificially increased because the two parts are segmented in the misaligned condition, and not segmented in the aligned condition (i.e., a methodological confound).

There are two typical counterarguments to the use of a physical gap between the top and the bottom half of a composite face. The first is that big composite effects are found without a gap. My answer is precisely that the effect could be artificially inflated, for the reason that I explained earlier. I don't believe that the conclusions from composite face studies that did not include a gap are erroneous because it remains a minor methodological problem. However, the fact remains that, in the studies that do not include a gap, the composite face effect might be artificially inflated. It is important to ensure that it is not the case, also considering that different populations compared (e.g., children vs. adults) can understand instructions about what is the top half of a face differently.

The second counterargument to the use of a gap is that it would introduce a distortion of the face structure, which might no longer fit our internal face template. This argument is not valid because faces are often perceived behind occluders in real life, and the perceiver might well consider such a gap as a small occluder. Also, if anything, it is often the *absence* of gap that plays a negative role for perceptual integration. Indeed, the human visual system tends to enclose a line or a space by completing a contour and ignoring such gaps in a figure, the so-called Gestaltist law of *closure* (Pomerantz & Kubovy, 1986; Wagemans, Elder, et al., 2012; Wertheimer, 1925/1967). As long as the gap is not so large as to break the continuity of the contour of the face, the visual system will readily complete the stimulus. However, when one does not include a gap, there is a contiguous border between the top and bottom halves. Consequently, the border, defined by a small variation of luminance and texture gradient, is enhanced (i.e., border contrast), so that subtle differences in luminance are more readily perceived (Mach, 1865; Ratliff, 1965). It is a case of figure–ground segregation, where borders are defined by luminance and also texture (Regan, 2000). Given that cells in V1 and V2 signal border ownership of luminance and texture contours (Chaudhuri & Albright, 1997; Nothdurft, Gallant, & Van Essen, 2000; Zhou, Friedman, & von der Heydt, 2000), without a gap, any slight difference in luminance and texture between the contiguous top and bottom halves is going to be enhanced (with a likely contribution of higher level visual areas as well; see Kastner, De Weerd, Desimone, & Ungerleider, 2000). Therefore, somewhat paradoxically, the face may appear as a more integrated and plausible combination of top and bottom halves when there is a gap than when there is no gap, what I'd like to call the *paradoxical gap composite illusion* (Figure 31).

For all these reasons, I recommend inserting a gap between the two face halves, unless participants unless another cue indicates the border between the two halves or unless another cue indicates the border between the two halves (e.g., a slight change of colour between the two halves; see de Heering et al., 2007). How big could/should the gap be? Ideally, it needs to be clearly visible and yet be as small as possible. The examples provided in this paper show that the composite visual illusion is compelling with physical gaps between the two halves. Note, however, that even with a relatively large gap, the composite face effect remains substantial (Taubert & Alais, 2009). This observation is interesting because it suggests that as long as upright faces are used, deviations from the experienced face morphology are tolerated to a certain extent (e.g., de Heering, Wallis, & Maurer, 2012; Taubert, 2009). In contrast, misaligning the two halves laterally even in the slightest way (about 8% of face width) disrupts completely the composite face effect (Laguesse & Rossion, 2011; see Figure 20).

Can object-based attention account for the composite face effect?. The composite face paradigm aims at measuring a perceptual

Figure 31. The paradoxical gap composite illusion. Looking at the two face stimuli on top, it is difficult, if not impossible, to tell which one is the real face and which one is a composite face (a face made of top halves and bottom halves belonging to different identities). Paradoxically, eliminating the physical gap between the two halves, as on the faces displayed below, makes this judgement trivial (at least on a computer screen): The face on the right is the composite face. Hence, having an actual gap inserted between the two halves of a composite face is not only advantageous to define objectively what "top half" the participants of a study should try to match, but it may in fact promote the perception of an integrated face stimulus rather than reducing it.

phenomenon. Yet, the effect obtained in a given experiment, especially at the individual level, could be due to many other factors than perceptual integration. One can exclude a decisional/response factor because the irrelevant bottom halves are associated with the same response ("different") in misaligned and aligned trials. Moreover, any putative *general* response bias between aligned and misaligned trials can be neutralized by including trials in which both halves are the same or both are different, in order to exclude such a bias (Figure 26). Also, as long as the top halves are presented at the same attended locations for aligned and misaligned trials, and that participants fixate the same spots in both cases (de Heering et al., 2008), there is no difference in overt spatial attention that could explain the results. However, misaligning facial halves breaks the face into two independent objects. Since ignoring a distractor that is located in a different object than a target is easier than if both are embedded in the same object (e.g., Kramer & Jacobson, 1991), one could argue that (covert) attention is reduced for the misaligned bottom half as compared to the aligned bottom half. More generally, because perceptual organization constrains attentional selectivity (e.g., Chen, 2012; Kimchi, 2009; Kramer & Jacobson, 1991), it may be argued that standard composite effect is due to a difference in (object-based) attention between aligned and misaligned trials. However, a putative difference in attention between aligned and misaligned faces does not necessarily mean that object-based attention *accounts* for the composite effect. Rather, in this situation at least, it is likely that perceptual integration (grouping) takes place *before* any attentional process, and could influence the *subsequent* allocation of attention (see Kimchi, 2009).

Independently of the complex issue of the relationship between perceptual grouping and attention, there are a number of observations that seems incompatible with an account of the standard composite face effect in terms of object-based attention. First, object-based attention theories would predict a substantial reduction of the composite face effect when a horizontal gap is included between the two parts, which is not the case (Figure 31; see earlier). Second, an object-based attention account is difficult to reconcile with larger composite effects for faces differing in shape than surface cues (Jiang et al., 2011), because in both cases the difference between aligned and misaligned trials, in terms of physical separation, is the same. Third, and more fundamentally for this issue, inversion offers an important additional control to misalignment because the stimulus remains a whole object. Yet, the composite illusion/effect disappears or is largely reduced for inverted faces (see earlier). Finally, the locus of the composite face effect in face-sensitive visual areas and on early visual ERPs with (Jacques & Rossion, 2009; Schiltz et al., 2010) or without (Kuefner, Jacques, Prieto, & Rossion, 2010; Schiltz & Rossion, 2006) concurrent behavioural responses, supports a perceptual locus of the effect independently or before any implication of putative attentional processes.

Other stimulus issues. For reasons explained earlier, although I do not encourage using greyscale rather than colour faces in face perception experiments because colour is a salient and diagnostic cue for face categorization, and of face identity in particular (Yip & Sinha, 2002), I advocate using greyscale faces in the behavioural composite face paradigm. Researchers should also try to minimize the abrupt variation of luminance between the top and bottom halves of the composite faces, which are present in many studies, and particularly enhanced when there is no gap between the two halves. The size of the composite stimuli is also an important issue:

They should not be too small so that fine-grained information to individualize faces can be extracted, but they should not be too big either, in order to see the whole of the face without having to make any eye movement. McKone (2009) addressed this issue for the holistic perception of *a* face, at the categorical level, and found that holistic processing peaked at distances functionally relevant for identification during approach (2–10 m; equivalent head size = 6–1.3 degrees). However, her experiments (e.g., perceiving a "Mooney" stimulus as a face) were not precisely aimed at testing identification, or individual face perception, so that it would be worth addressing this issue with the composite face paradigm.

Another important issue is the homogeneity of the face set that is used to create the composite faces. Although it is important to maintain the natural variations in face shape, one should avoid creating composite faces across sex (e.g., a female top face with a male bottom face), or age (e.g., a young top face with a bottom old face), or "race" (e.g., an Asian top face with a bottom Caucasian face). Otherwise, the influence of the bottom half on the top half could be due to a perceived change in age, sex, or "race" rather than identity, an obvious confound in the paradigm. As a matter of fact, there is evidence that sex and age judgements of composites' top halves are biased towards the bottom halves' sex and chronological ages, respectively (Baudouin & Humphreys, 2006; Hole & George, 2011), the effect of age being also due to face shape rather than surface cues.

In the same vein, only faces with a neutral—or constant—expression should be used to test holistic processing of facial identity. If a slight smile is present on the bottom half of a face, it is not only the identity of the top half that changes but also its expression: There is a persistent illusion that the eyes have a "smiling" expression (Figure 32). This observation is in line with studies showing robust composite face effects for judgements of expression (Calder, Young, Keane, & Dean, 2000; Palermo et al., 2011; Tanaka, Kaiser, Butler, & Le Grand, 2012). In fact, since facial halves are considered less trustworthy (attractive) when paired with untrustworthy (unattractive) rather than trustworthy (attractive) halves (Abbas & Duchaine, 2008; Todorov, Loehr, & Oosterhof, 2010), faces should or could even be paired for equal levels of attractiveness if one wants to avoid the possibility that judgements of identity are confounded by perception of changes in trustworthiness/ attractiveness.

Summary and conclusions of Part 2

The composite face illusion is a compelling visual illusion in which changing the bottom halves of faces makes the whole faces, including the unchanged top halves, look different (Figure 1). Inserted in a behavioural same/different matching task, this illusion leads to a simple paradigm that has been used in

Figure 32. The smiling composite face illusion. One cannot help but see a "smiling expression" in the region of the eyes, even though the top half of the face has a completely neutral expression (just hide the bottom half, below the white line).

more than 60 studies so far to inform about perceptual integration, or grouping, of facial parts into a whole face. Several other paradigms have been used to measure holistic/configural face perception, or perceptual integration. For instance, Sergent (1984) used a matching task with faces varying in one, two, or three "configural" or "featural" manipulations, and showed that these manipulations were not processed independently from one another (see also Amishav & Kimchi, 2010; Barton, Zhao, & Keenan, 2003). Tanaka and Farah (1993, 2003; see also Davidoff & Donnelly, 1990; Donnelly & Davidoff, 1999) developed the whole–part advantage paradigm with faces, a paradigm which is used to study perceptual grouping of simple elements in two-dimensional shapes (Pomerantz & Portillo, 2011; Pomerantz, Sager, & Stoever, 1977). In Tanaka and Farah's (1993) paradigm, the discrimination of two isolated facial parts is enhanced by the addition of a whole facial context (see Figure 33A), an effect that disappears for inverted faces. Another paradigm derives from the so-called Thatcher illusion (Thompson, 1980, 2009), in which an upright but not an inverted face appears grotesque if its parts are inverted (Figure 33B). The Thatcher paradigm has also been used in a number of behavioural studies to investigate the interdependence of facial parts (e.g., Lewis & Johnston, 1997; Murray et al., 2000; Rhodes et al., 1993; Stürzel & Spillmann, 2002). As mentioned earlier, behavioural studies using these other paradigms have provided information about holistic/configural face perception that generally agrees with studies using the composite face paradigm.

Nevertheless, it is not by chance that the composite face paradigm is the most widely used in studies of holistic face perception. Here, I would like to list a few of the (good) reasons why I think this paradigm is so popular among researchers in the field.

1. The paradigm is based on a *visual illusion*, so that one can appreciate visually what is meant by holistic/configural face perception. Other paradigms also derive from visual illusions, or can be illustrated as visual illusions (Figure 33). However, the composite face illusion is almost as compelling as the Thatcher illusion, and the composite face

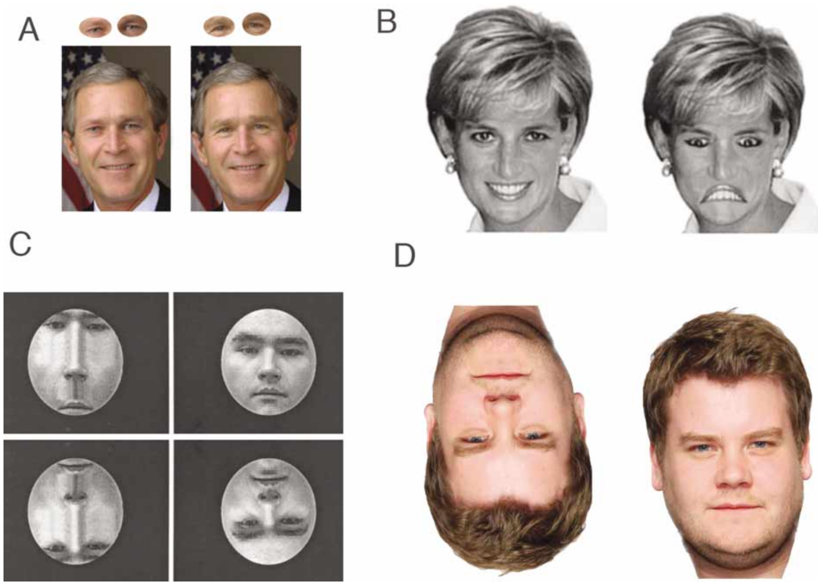

Figure 33. A few visual illusions that appear to reflect holistic/configural face perception. (A) The whole–part face illusion (courtesy of Jim Tanaka), in which it is easier to perceive who is the familiar face when the eyes are inserted in the whole face picture than when they are presented in isolation (Tanaka & Farah, 1993). It is an effect of context (as used with simple object shapes, e.g., Pomerantz & Portillo, 2011). Official photograph of President George W. Bush.
Source: Executive Office of the President of the United States. Photo by Eric Draper, White House (2003). (B) The "Thatcher" illusion (Carbon, 2007), in which the eyes and mouth are inverted but the face is perceived as grotesque only when it is upright, not when it is inverted. (C) An illusion created by Lee and Freire (1999) in which the shape appears oval rather than round when the face features have been expanded, an illusion that (almost) disappears when the face is presented upside-down. (D) The "Fat Face Thin Illusion" (Thompson, 2012) in which an inverted face appears thinner than its upright version, probably because the width of the bottom half of the face influences the perception of the top when it is upright, not when it is inverted (see also Sun et al., 2012). To view this figure in colour, please see the online issue of the Journal.

paradigm has many additional strengths compared to a paradigm based on the Thatcher illusion.

2. The composite face paradigm is a *simple* paradigm, having usually only two conditions of interest: Aligned versus misaligned faces for "same" trials.
3. The control condition, spatial misalignment of the two face halves, is *theoretically grounded* in Gestalt Psychology (see earlier).
4. The definition a "part" is *objective* in the paradigm, at least when a gap is included. It means that participants know exactly what they have to consider as a "part" in the task. In contrast, when using whole faces, asking participants to match "the eyes" or the "Mouth + Nose" (e.g., Goffaux, 2009) is much more ambiguous (to which part of the face does the "Mouth + Nose" exactly refer?).
5. The manipulation for the control condition is *objective*: A horizontal cut of the face in two parts. Of course, the alignment manipulation can vary across studies (height of separation between the two halves, a gap or no gap included, the size of the gap, the amount of spatial alignment, etc.). However, these methodological issues may not be critical and can, at least, be defined objectively (i.e., quantified). In contrast, the Thatcher illusion and whole–part paradigm depend a lot on which facial parts are manipulated (eyes or mouth) and how they are defined (e.g., eyes including eyebrows or not?). The Thatcher illusion is also highly dependent on facial expression: It works very well if the face is expressive, with the mouth wide open, but not so well if the mouth is closed as in a neutral expression.
6. The spatial misalignment is *easy* to do in the composite face paradigm. It can be applied on full-face photographs, without requiring sophisticated graphics skills. In contrast, pasting the eyes or mouth of a face onto another face in the whole–part paradigm, or moving the facial parts horizontally or vertically in the face in so-called "configural" manipulations, are not simple. In such studies, the quality of stimuli differs depending on the experimenter's skills and care, and schematic faces are sometimes used to facilitate these manipulations (e.g., Tanaka & Farah, 1993; Tanaka & Sengco, 1997). Contrary to the Thatcherized faces, the composite faces are still real faces, and can even look almost as veridical as original faces if they are made carefully (with the exception of the gap in between the two halves).
7. There are *no issues of manipulations* of "configural" versus "featural" cues in the composite paradigm, unlike in the whole–part advantage paradigm, which makes the composite paradigm less open to misinterpretation of the nature of the effects.
8. The *variables measured* in the composite face paradigm are *objective* (accuracy and RTs in a matching task), rather than subjective as in the

Thatcher paradigm. In the latter paradigm, participants are usually asked to subjectively judge the grotesqueness of faces presented upright or upside-down, which is not ideal for quantitative measurements and prevents the use of RTs as a variable in the task.

9. This objectivity makes the composite face paradigm straightforward to test with approaches measuring *neurophysiological responses* in face identity adaptation paradigms (Jacques & Rossion, 2009; Schiltz & Rossion, 2006). In contrast, it is more difficult to interpret differential responses, or different neural adaptation effects, to a veridical and a Thatcherized face (e.g., Boutsen, Humphreys, Praamstra, & Warbrick, 2006; Carbon, Schweinberger, Kaufmann, & Leder, 2005).
10. The possibility to use (mis)alignment as a control makes the composite paradigm *independent of inversion*, unlike paradigms based on the Thatcher illusion. As much as face inversion is a great manipulation, it is a control condition that has its limitations (a change of the location of the fixation point, and of the respective amount of visual stimulation in the upper and lower visual fields). It is a good thing to have the possibility of using two types of control manipulations (misalignment and inversion), and the composite face paradigm offers that.
11. The composite face paradigm gives rise to *large effects* (sometimes around 20% for accuracy rates or RTs, sometimes on both variables, e.g., Rossion & Boremanse, 2008), and if the paradigm is applied properly it is rare that an individual participant in a given experiment does not show an effect. The whole–part advantage paradigm does not give rise to such large effects, which makes it much more difficult to study at the individual level, for instance when comparing a single case of prosopagnosia to normal controls (see Ramon et al., 2010).
12. The composite face paradigm leads to effects that are highly *specific*: They are not found for nonface object shapes (Gauthier et al., 1998; Macchi Cassia et al., 2009; Robbins & McKone, 2007; Taubert, 2009). This is not always the case for the whole–part paradigm (e.g., Donnelly & Davidoff, 1999; Seitz, 2002), something that might have to do with how the parts are defined in this latter paradigm, or the fact that the whole–part paradigm captures a more general measure of context. Yet, because it involves an objective manipulation (spatial misalignment), the composite face paradigm can be applied to nonface object shapes relatively easily.

All these advantages probably account for the popularity of the composite face paradigm, and strengthen the findings made with this paradigm, reviewed here. Nevertheless, the composite face paradigm also has some limitations that should be mentioned.

The main weaknesses of the paradigm come from the fact that it provides a behavioural measure. Thus, even if it has been designed to isolate holistic face perception by comparing two conditions differing only by one factor (usually misalignment), there are many processes taking place between sensory perception and motor output that can contribute to the performance in each of the conditions (perception, selective/spatial attention, working memory, decisional factors, etc.). These processes can create spurious differences or, contrarily, cancel out any difference between these two conditions. In a typical group study, isolating the signal from these sources of noise is done by averaging data across trials and participants, so that the nondesirable contribution of these processes, i.e., the noise, cancels out. Thus, in a group design, one can hope to isolate the signal and observe a difference between two or more conditions in the magnitude of the composite face effect. However, because of the measure is noisy, it might be very difficult to estimate the magnitude of holistic perception at the level of a single participant with such a paradigm, unless one uses many trials and several testing sessions. For this reason, but also because there is a large amount of variance across participants in terms of the magnitude of the effect, with some normal participants of a given study not always showing a significant effect, one should remain careful in concluding for a lack—or even the normality—of holistic face perception in a single case of prosopagnosia tested only once for instance (e.g., Ramon et al., 2010; Rezlescu, Pitcher, & Duchaine, 2012). For the same reason, correlation measures based on interindividual variance at the composite face paradigm may not be high or even significant, as discussed earlier. Moreover, the large amount of variance across participants in terms of the magnitude of the effect may sometimes prevent disclosing significant differences, even when the magnitude of the composite face effect appears to be reduced substantially (for instance, between normal and contrast-reversed faces in Figure 2 of the study of Taubert & Alais, 2011). Finally, because the outcome of the composite face paradigm is a behavioural measure, its functional locus is unknown, and one has to rely on other approaches, such as ERPs and fMRI to gain information about that. These issues are worth being reminded to avoid mis- or overinterpretations from the presence or absence of composite face effects. However, they are general issues of cognitive psychology, concerning all behavioural measures obtained in such paradigms.

Beyond these general issues, specific weaknesses of the composite face paradigm should also be mentioned. One limitation is that spatial misalignment of the top and bottom halves of a face is a control that affects low-level processes, not only high-level processes. Inversion also suffers from this limitation, because it entails a change of fixation in the paradigm (from the upper to the lower visual field). Thus, the two manipulations are somewhat complementary and, ideally, both should be used to ensure that a new effect observed in the paradigm is really solid. Another limitation is that the effect is found either in accuracy, or RTs, or both, so that it is necessary to consider both variables, and it is difficult to compare the magnitude of the effect across participants (which variable should be given more weight than the other?). Moreover, in my experience, participants usually realize during the course of a matching experiment that some of the top halves look slightly different but should probably be nevertheless associated with a "same" response, given that in other trials (the "real different" trials) the perceived difference is more salient. Thus, an effect observed in accuracy at the beginning of an experiment might shift towards an effect observed in RTs at the end of the experiment. More generally, the fact that the effect is usually distributed between two variables makes it difficult to use correlation measures based on the magnitude of the composite face effect across individuals. Finally, some weaknesses have more to do with the current status of the literature on the composite face paradigm. Personally, I would be keen on seeing more data on normal and contrast-reversed faces because the available data (e.g., Figure 2 of Taubert & Alais, 2011)—and the composite illusion (Figure 12)—suggest a reduction of the composite face effect following contrast reversal (albeit less than following inversion). It also seems to me that the abnormal vertical displacement of the eyes (i.e., distortion of face configuration) does reduce the composite face effect (de Heering, Wallis, & Maurer, 2012) qualifying the conclusions of this recent study. I picked these two studies because they drew conclusions from an absence of evidence (i.e., a significant difference) despite observing nonsignificant differences in the predicted direction. Such effects might have been revealed by computing inverse efficiency, or increasing the power of the experiment. Interestingly, in the recent study of de Heering, Wallis, & Maurer, (2012), the composite effect was reduced for typical faces when they were mixed up with distorted faces (Exp. 2) compared to when they were presented in block (Exp. 1). This suggests that the face template may not be stable when distorted and undistorted faces are mixed up in the design. This blocking factor could be systematically manipulated in composite face future studies.

Despite these caveats, admittedly, these findings and other observations suggest that, as long as upright faces are considered, the degree of generalization of the composite face effect to facial morphologies that are not visually experienced needs more clarification. For instance, Mondloch et al. (2010) did not replicate the larger composite effect for "same-race" than "other-race" faces (Michel, Rossion, et al., 2006), possibly because Caucasian participants of the Mondloch et al. study did not present with any "other-race face effect" in recognition performance and were exposed to Asian faces through their national media (see also Rossion & Michel, 2011, for a discussion of this issue). The equally large composite effect for human and chimpanzee adult faces (but not other species such as monkeys, sheep, or birds) is undisputable in terms of data (Taubert, 2009). Yet, this is also a surprising result, considering that the composite effect is reduced even for "other-age" or "other-race" faces within the human species, in several studies (see earlier).

Another issue is that the specificity of the effect for faces has been tested (and established) only against a few categories of nonface stimuli (profile dogs in Robbins & McKone, 2007; cars in Macchi-Cassia et al., 2009; "Greebles" in Gauthier et al., 1998; "sticks" in Taubert, 2009; body shapes in Soria-Bauser, Suchan, & Daum, 2011, although see Robbins & Coltheart, 2012, who found a weak composite effect for both upright and inverted body shapes), and should be strengthened (or qualified) by further tests.

To resolve these issues, we certainly need more tests of the magnitude of the composite face effect under systematic (i.e., parametric) manipulations of face stimuli (e.g., small to large, undistorted to completely distorted faces, etc.) and design (relative proportions of "same" and "different" trials, duration of presentation, etc.).

Finally, a substantial composite face effect when focusing on the bottom half of faces (e.g., Nishimura et al., 2008) is puzzling because of the near absence of composite face illusion for bottom halves of faces (Figure 29). This is a concern that needs to be addressed, and experimenters first need to ensure that human observers are able to keep eye gaze fixations on the bottom halves of the faces, equally for aligned and misaligned faces, when instructed to do so.

Before moving to Part 3, I would like to make a couple of additional points. First, almost implicitly, by using the composite paradigm, researchers have realized that there is a strong asymmetry between the top and bottom halves of a face, and between the judgements of identity ("same face") and of different identities ("different face"). These are important aspects of face perception, and it seems that holistic face perception is particularly important when having to judge that faces are the *same* from the *top* half of the face.

Second, it is worth noting that in the composite face matching paradigm, the *exact same image* is used between the two top halves to match (for an exception, see Hole et al., 1999; consider also the change of image size in some studies). Yet, despite the fact that observers could, in principle, rely on a simple image-based matching strategy, they apparently cannot do it in practice. They make mistakes and take more time to judge the identity of two strictly identical top face images because different images are present at the bottom halves. This observation shows that in the composite face paradigm observers cannot use a simple image-based matching strategy on one part of a face without being influenced by the other part(s), and that despite some criticisms (e.g., Megreya & Burton, 2006), the matching of identical images of unfamiliar faces *can* be highly relevant to understand face perception.

Finally, the composite face paradigm measures *perceptual integration*, or grouping, of two parts into a whole face. Two new whole face identities are created in the aligned version, and they have to be compared to make a "same" judgement. Hence, in this particular paradigm, performance drops in the condition associated with holistic face perception as compared to the condition that is not associated with holistic face perception. However, despite this *negative* effect of holistic face perception, it is *not* a paradigm that aims at measuring *interference*, or a negative influence of one face part (the bottom part) on another part (the top part), as it is sometimes described. If one refers to interference in the paradigm, it is because perceptual integration (of the target and distracter halves) interferes with performance on the target half. Thus, the interference comes from the comparison of the two faces and the nature of the task, not from the intrinsic relationship between the two face halves. This issue is extremely important to ensure that the composite face paradigm is not confused with a different approach of the problem inspired from an attentional framework in experimental psychology, and which will be the topic of the third and last part of this review.

PART 3: THE ILLUSION OF A MEASURE

The congruency/interference paradigm with composite faces

In the last decade, a group of authors led by Isabel Gauthier and Jennifer Richler have developed a different kind of paradigm that uses composite faces in a delayed matching task: The *congruency/interference composite face paradigm*. Based on their results with this paradigm, these authors have made a number of claims concerning the nature (functional locus, specificity, etc.) of holistic face processing (Bukach, Bub, Gauthier, & Tarr, 2006; Cheung, Richler, Palmeri, & Gauthier, 2008; Richler, Cheung, & Gauthier, 2011a, 2011b; Richler, Cheung, Wong, & Gauthier, 2009; Richler, Gauthier, Wenger, & Palmeri, 2008; Richler, Mack, Gauthier, & Palmeri, 2009; Richler, Tanaka, Brown, & Gauthier, 2008; Richler, Wong, & Gauthier, 2011). In addition, Gauthier, Richler, and colleagues (GRC) have claimed that the standard composite face paradigm reviewed in Parts 1 and 2 was a "partial" version of their congruency paradigm, the latter being "complete" or "full" (Cheung et al., 2008; Gauthier & Bukach, 2007; Richler, Cheung, & Gauthier, 2011a, 2011b; Richler, Mack, et al., 2011). GRC also stated that the standard composite face paradigm had "poor construct validity" (Richler, Cheung, & Gauthier, 2011a, p. 1) and "poor reliability" (Richler, Cheung, & Gauthier, 2011b, p. 467), and was essentially "flawed" (Richler, Cheung, & Gauthier, 2011b; Richler, Mack, et al., 2011). Or, to quote these authors, that "we cannot hope to make theoretical progress in our understanding of the mechanisms underlying face perception if we continue to use the partial design of the composite task" (Richler, Cheung, & Gauthier, 2011b, p. 470). That is, according to these authors, studies that have used the standard composite paradigm to understand holistic face perception have led to incorrect conclusions, so that it should be abandoned and the studies redone with these authors' own congruency/interference paradigm.

GRC's constant criticism of the standard composite face paradigm seems unjustified. By proposing that the standard composite paradigm is abandoned, these authors take the risk of leading the field of face processing in the wrong direction. There is also a risk of dismissing a substantial amount of past findings that are relevant for our understanding of the nature of holistic face perception. Contrary to these authors, I believe that the standard composite face paradigm has served, and can serve, the field of face processing very well. It can certainly be improved and needs to be complemented by other measures, as I discussed at the end of Part 2, but it deserves much better than such a flat dismissal.

In Part 2 of the present paper, I explained the rationale behind the standard composite face paradigm, in which two conditions differing only by a single factor (spatial alignment) are compared. Thus, it is a methodologically sound paradigm and in this last part of the paper, I will not spend more time dispelling the myth that the standard composite face paradigm is methodologically flawed. Rather, I will argue that the congruency/interference paradigm used by GRC with composite faces belongs to a general class of paradigms that are not aimed at measuring perceptual integration but which have rather been used in experimental psychology since Stroop (1935) to measure attentional interference and response conflict effects. Then I will argue that if one wants to interpret its outcome in terms of perceptual integration ("holistic processing"), the congruency/interference composite face paradigm presents three limitations. First, the irrelevant face half is associated with a behavioural response that either conflicts or agrees with the response associated with the target face half. Second, the paradigm lacks a control condition (misaligned or inverted trials). Third, it gives the same weight to "same" and "different" trials in the computation of the effect, although only "same" trials should be considered if one aims at measuring holistic perception. Next, I will show that GRC's particular version of the congruency paradigm contains numerous additional attentional and stimulus confounds. Finally, because of these confounds, studies that have used GRC's congruency/interference composite paradigm have led to observations that are almost impossible to interpret and cannot be related to holistic face processing in an intelligible way.

I would like to warn the reader that in Part 3 of this review, especially later, I will be sometimes quite critical towards the congruency/interference composite paradigm developed by a particular group of authors, and the results that these authors obtained with this paradigm. I apologize to these authors and to the reader for that, but I am convinced that such a critical analysis is necessary at this stage and will give us more solid ground on which to move forward with our understanding of holistic face perception.

The roots of the congruency/interference composite face paradigm. GRC's rationale for designing the congruency/interference composite face paradigm (Figure 34) is not based on a visual illusion, and the paradigm does not aim at capturing a perceptual phenomenon. Rather, the theoretical framework behind that paradigm is attention and interference, along the lines of the famous Stroop paradigm (MacLeod, 1991; Stroop, 1935). Indeed, GRC define holistic face processing primarily in attentional terms, namely as "a failure of selective attention" or "the failure of selective attention to face parts" (e.g., Richler, Gauthier, et al., 2008, p. 332; Richler, Tanaka, et al., 2008, p. 1357; Richler, Cheung, & Gauthier, 2011b, suppl. material, p. 22; Richler, Wong, & Gauthier, 2011, p. 130), or "the inability to selectively attend to one (face) part while ignoring information in another part" (e.g., Richler, Tanaka, et al., 2008, abstract). Other authors also noted the parallel between GRC's paradigm and the Stroop paradigm (Robbins & McKone, 2007), and GRC themselves refer to the attentional literature in general in their studies, and often explicitly to the Stroop paradigm (e.g., Richler, Mack, et al., 2009, 2001; Richler, Cheung, et al., 2009).

This conceptualization of holistic (face) processing is quite different than what we discussed so far. In the composite illusion, one *sees* two identical top halves of faces as being different when they are aligned with different bottom halves. It is a perceptual phenomenon, which is thought to reflect perceptual integration, and the composite face paradigm attempts to capture it. This phenomenon is not conceptualized as a failure of selective attention: During the composite face task, an observer is asked to judge the top half of a face only and keep fixation on it (de Heering et al., 2008). Although one cannot exclude the possibility that object-based attentional factors are involved in the standard composite face effect of misalignment (see earlier), there is no evidence suggesting that this top half is less (overtly) attended to when it is aligned rather than misaligned with its bottom half. As explained at the end of Part 1 of this review, this phenomenon should not be attributed to a kind of interference between the top and bottom face halves either. However, if one is interested in measuring primarily *attentional interference* rather than *perceptual integration* between face parts, then it makes sense to use a fundamentally different paradigm than the standard composite face paradigm.

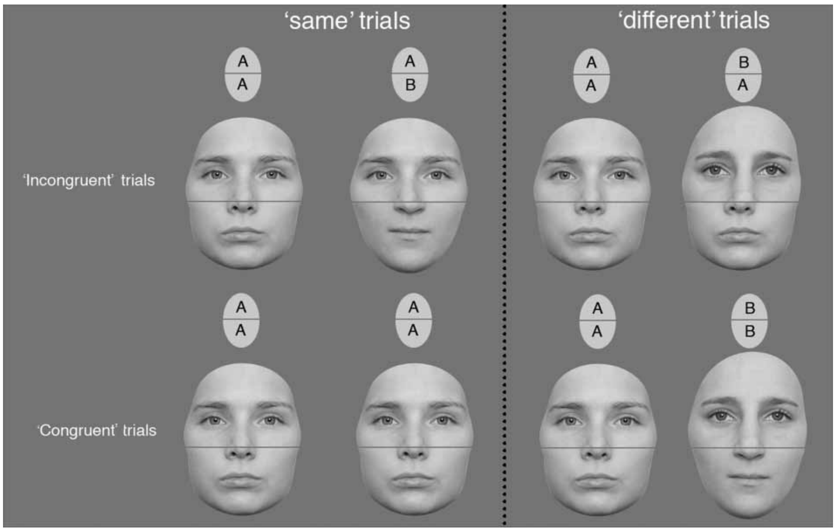

Figure 34. The congruency/interference paradigm of Gauthier, Richler, and colleagues. Here, only top face halves are considered. In incongruent trials, the bottom face halves do not lead to the same behavioural response as the top face halves. The authors consider a lower performance for all incongruent trials as compared to congruent trials as reflecting "holistic face processing".

In a paradigm aiming at measuring attentional interference with visual stimuli, the observer is asked to attend to one aspect of a visual display (the target), while there is an irrelevant aspect (the context) that is there to interfere. In a typical Stroop colour–word paradigm, for instance, one should attend to and read the ink colour (target) of a printed word (context). Stroop (1935) showed that naming the colour of an incongruent colour word (i.e., saying "blue" when presented with the stimulus "red", coloured blue) was slower than the naming of the colour presented on a small solid square ("v", coloured blue). This paradigm was later extended to compare a condition in which the word is incongruent with the colour (saying "blue" when presented with the stimulus "red", coloured blue) to a condition in which the word is congruent with the colour (saying "blue" when presented with the stimulus "blue", coloured blue), respectively (Dalrymple-Alford & Budayr, 1966; Sichel & Chandler, 1969; see Figure 35A). Participants of these experiments usually respond faster in the congruent than in the incongruent condition (see MacLeod, 1991).

A congruency/interference paradigm can also be used in a delayed matching task, for instance to decide if two consecutive colours are identical: "3" [coloured blue] then "3" [coloured blue] ("same" response expected, the response based on the target—the colour—being congruent with the response based on the context—the number) or "3" [coloured blue] then

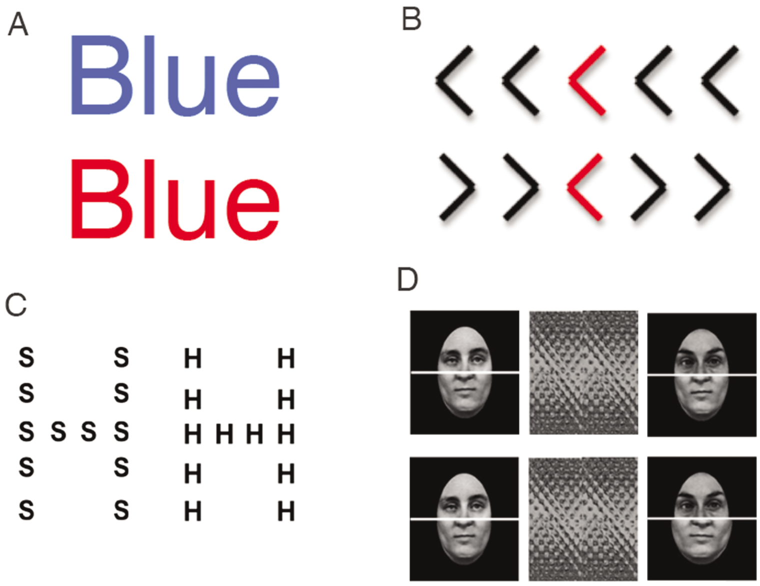

Figure 35. A family of paradigms measuring attentional interference and response conflict. (A) A variant of the Stroop design, introduced originally by Stroop (1935) and extended to congruent/ incongruent paradigms later on. Naming the colour of the stimulus is slower for the item printed in the colour that does not correspond to the word. (B) A variant of the Eriksen flanker task (Eriksen & Eriksen, 1974) with congruent (on top) and incongruent (below) flankers surrounding a central target. (C) Navon (1977) types of stimuli, in which a large letter is composed of small letters that are either incongruent (on the left) or congruent (on the right). (D) The congruency paradigm of Gauthier et al., in which different top halves of faces are associated either with identical bottom halves ("incongruent trials", on top) or different bottom halves ("congruent bottom halves", below). The paradigms illustrated in A, B, and C are used to test for attentional interference between different representations, the dominance of one onto the other, or response conflict monitoring; but not to test the perceptual integration of these representations. To view this figure in colour, please see the online issue of the Journal.

"4" [coloured blue] ("same", incongruent). In such a paradigm, there are also trials in which the response should be "different": "3" [coloured blue] then "4" [coloured red] ("different", congruent) or "3" [coloured blue] then "3" [coloured red] ("different", incongruent). In sum, there are trials in which both the target and the context remains the same, or both differ, between the two visual displays (congruent trials). And there are trials in which either the target *or* the context changes (incongruent trials).

Replacing the number by the top half and the colour by the bottom half face gives a congruency/interference composite face paradigm. More precisely, if one applies this modified Stroop paradigm to two halves of a face, with the top halves as targets, "same" trials are those trials in which the top half face is associated with a correct "same" responses: "Top A/bottom

A" to "top A/bottom A" ("same", *congruent*), or "top A/bottom A" to "top A/bottom B" ("same", incongruent). And there are trials for which a correct response is "different": "Top A/bottom A" to "top B/bottom B" ("different", congruent) and "top A/bottom A" to "top B/bottom A" ("different", incongruent). This is the kind of congruency/interference paradigm used by GRC with faces (Bukach et al., 2006; Cheung et al., 2008; Richler, Cheung, & Gauthier, 2011a, 2011b; Richler, Gauthier, et al., 2008; Richler, Mack, et al., 2008, 2009, 2011; Richler, Wong, & Gauthier, 2011) and nonface objects (e.g., Bukach, Phillips, & Gauthier, 2010; Gauthier & Tarr, 2002; Wong et al., 2011, 2012; Wong, Palmeri, & Gauthier, 2009). In all of these studies, better performance for congruent as compared to incongruent trials is taken as evidence for a "failure of selective attention", which is conceptualized as reflecting "holistic face processing". In short, Gauthier and colleagues should be credited for having extended the ubiquitous congruency/interference paradigms used in experimental psychology, most notably the Stroop paradigm, to composite faces[4] (Figure 34).

Note that in the congruency/interference paradigm one cannot determine whether it is the incongruent context face half that *interferes* with the processing of the target face half, or if it is the congruent context face half that *facilitates* the processing of the target face half. To clarify this classical issue in the attentional interference literature (MacLeod, 1991), one would need to include a neutral condition, for instance a target half presented in isolation (see, e.g., Goffaux, 2012, for isolated eyes). Interestingly, this is what Stroop (1935) did in his original study: He showed that naming the colour of an incongruent colour word (i.e., saying "blue" when presented with the stimulus "red" [coloured blue]) was slower relative to the naming of the colour presented on a small solid square ("v" [coloured blue]), a neutral condition. There was no colour word in congruent ink colours in Stroop's original study, just as congruency is not manipulated in the standard composite face paradigm. Based on this, GRC would have certainly characterized Stroop's original paradigm as being partial, and flawed.

Note also another difference between the Stroop (1935) paradigm and the congruency/interference composite face paradigm. In the original Stroop paradigm, the two dimensions of the stimulus that interfere with each other—namely the colour and the name of the word—spatially overlap (for exceptions see Wühr & Waszak, 2003), unlike the two face halves in the congruency/interference composite face paradigm. In this respect, this latter paradigm more resembles other congruency/interference paradigms analogous to the Stroop task, such as the "flanker" task (Eriksen & Eriksen,

[4] Analogues of the Stroop type of paradigm have been used before in face research (de Haan, Young, & Newcombe, 1987; Young, Ellis, Flude, McWeeny, & Hay, 1986), but they take a very different form from the composite task.

1974). In this task, a directional response (generally left or right) is made to a central target stimulus. The target is flanked by nontarget stimuli which correspond either to the same directional response as the target (congruent flankers) or to the opposite response (incongruent flankers) (Figure 35B). It is generally found that response times are slower for incongruent stimuli than for congruent stimuli. This paradigm was developed to assess the ability to suppress responses that are inappropriate in a particular context, and the slower responses seem to be due to an attentional-selection problem (MacLeod, 1991) or a conflict at the level of the response (Eriksen, 1995).

Another famous interference paradigm is the Navon task (Navon, 1977, 2003), in which a large letter ("H") is composed of small letters that can be *congruent* ("hh") or *incongruent* ("ss"). When required to respond to the large letter, the congruency of the small letters have much less impact than when required to respond to the small letters (Figure 35C).

Importantly, this latter paradigm, like the paradigms described previously, is not designed to investigate perceptual integration between large and small letters. Rather, it uses interference to assess whether the large letter or the small letters dominate the process.[5] More generally, in these paradigms, whether the critical factor is the size of the letters used (Navon), the word and ink colour (Stroop), or the central and peripheral arrows (Eriksen), the effects are interpreted in terms of attentional interference and/or response selection conflicts (Eriksen, 1995; Hommel, 1997; MacLeod, 1991; see also Goldfarb & Henik, 2006) or more rarely in terms of competition between different semantic representations (Luo, 1999). GRC's congruency/interference paradigm with face halves has been developed using the exact same logic (Figure 35D), and thus it is not surprising that the effects obtained with this paradigm—"holistic processing" according to the authors—are often interpreted in terms of attentional and decisional processes rather than perception by the authors (see later). In fact, it is very likely that such factors contribute heavily to the effects obtained in this congruency/interference composite face paradigm. However, it does not mean that these factors have something to do with holistic processing as defined in terms of perceptual integration, and play a role in the effects obtained in the standard composite face paradigm described in Parts 1 and 2.

The congruency paradigm has a built-in response conflict confound. At a general level, the standard composite face effect can be

[5] A global precedence effect in the Navon task is sometimes interpreted as reflecting "holistic processing", although this does not imply that the small letters (components) integrate better to form the large letter when they are congruent with its identity (see Kimchi, 1992, for a discussion of this issue).

characterized as an effect of *context*: It is based on the comparison of two conditions in which the context is either aligned or misaligned with the target. The standard whole–part advantage paradigm (Tanaka & Farah, 1993; see Figure 33) is also a paradigm that is based on context, this time on the *presence* or *absence* of a context. Alternatively, it is sometimes based on the normal configuration of this context (Homa, Haver, & Schwartz, 1976; Mermelstein, Banks, & Prinzmetal, 1979; see also Gorea & Julesz, 1990; Suzuki & Cavanagh, 1995). Following the seminal work of Garner (1974), some studies have also used a context and a target to make inferences about perceptual grouping (Pomerantz, Carson, & Feldman, 1994; Pomerantz & Pristach, 1989; Pomerantz, Pristach, & Carson, 1989), and this latter approach has been applied to faces (Amishav & Kimchi, 2010; Pomerantz et al., 2003). Such contextual effects, in these paradigms, are generally used to make inferences about holistic perception, in the sense of perceptual integration or perceptual grouping. Garner's (1974) paradigm is even defined as measuring the *interference* from a context on a target. However, critically, in all these paradigms, the context is *never* associated with a dual behavioural response that conflicts *or* agrees with the dual behavioural response of the target.

In the standard composite face paradigm, the context is aligned or misaligned with the target. However, in the relevant ("same") trials, the behavioural response associated with the context is exactly the same in both cases (always a "different" response; thus, always "incongruent" with respect to the response associated with the target). That is, if the contexts (bottom halves) were presented alone, there would be only one kind of correct response ("different"), for both conditions. In the whole–part advantage paradigm, the context is present versus absent (Tanaka & Farah, 1993; see Figure 33). When it is present, it is neutral because it is exactly identical in the two alternatives of this forced choice matching task (Tanaka & Farah, 1993). Thus, if the two contexts were presented without the targets in the whole–part advantage paradigm, the participant would be unable to choose which alternative is correct. In the Garner paradigm, interference is considered when there is an increase in RT and/or error rates to a target (one element of a display) caused by random trial-to-trial variation in the context (another, irrelevant, element). However, importantly, it is the *variability* of the context that matters and *not* its response *incongruency* with the target (Pomerantz et al., 2003). That is, the context is not associated with a dual behavioural response that could agree or disagree with the dual behavioural response associated with the target, excluding an account of the effect at the response level (Garner, 1988). In summary, all these paradigms are implemented so that a response based on the context alone is completely neutral: Either it is impossible to make, or it leads to the exact same behavioural response in the two conditions compared.

In contrast, in the congruency/interference composite face paradigm of Gauthier, Richler, and their colleagues, and also in a few paradigms used by other authors in which different parts or cues of faces are manipulated (e.g., Anaki, Nica, & Moscovitch, 2011; Farah et al., 1998; Goffaux, 2009, 2012; Meinhardt-Injac, Persike, & Meinhardt, 2010, 2011), the context is not behaviourally neutral. That is, in the conditions that are compared, the context is associated with a behavioural response that either agrees or conflicts with the behavioural response of the target (Figure 36). In reality, all these studies follow a rather unfortunate modification of the whole–part paradigm introduced by Farah et al. (1998, Exp. 1) to a same/different judgement, in which the target (relevant) part of a face is associated with a dual behavioural response, and the context (irrelevant parts) with the same dual behavioural response.

This modification makes the interpretation of an effect of context in such modified paradigms highly ambiguous. That is, the target and the context could be processed completely independently, in parallel, but in the incongruent condition they are associated with conflicting behavioural outputs, and in the congruent condition they are associated with the same output. Thus, a difference between congruent and incongruent conditions could be entirely due to a conflict occurring at the response level. This is not just hypothetical. As a matter of fact, in similar paradigms such as the Eriksen flanker task, interference from a perceptually incongruent flanker is drastically decreased, or even eliminated, if flanker and target are mapped onto the same rather than conflicting responses (Eriksen & Eriksen, 1974; Miller, 1991), implying that it is not visual dissimilarity per se that is responsible for the typical flanker effect, but the fact that flanker and target lead to alternative, conflicting responses (Eriksen, 1995; Hommel, 1997). In the same vein, if Garner's paradigm is used in a same/different matching task, in which the target and the distractor leads to a response conflict, dimensions that are separable on the typical classification task such as colour and shape, become "integral" just because of a conflict at the response level (Garner, 1988).

Given this, it will not come as a surprise that Gauthier, Richler, and colleagues sometimes interpret the effects obtained in their congruency paradigms with composite faces in terms of a decisional process[6] (see later).

[6] However, the authors seem to distinguish decisional processes from response conflicts, even excluding response conflicts such as accounting for holistic processing based on the outcome of a naming task with learned composite faces (Richler, Cheung, et al., 2009). Interestingly, the results of this latter study do not fully support the authors' argument (Figure 1 of that paper) and suggest a contribution of a naming response conflict. In any case, this study does not exclude at all that any effect observed in the same/different matching version of the congruency/interference paradigm could well be due to a response conflict.

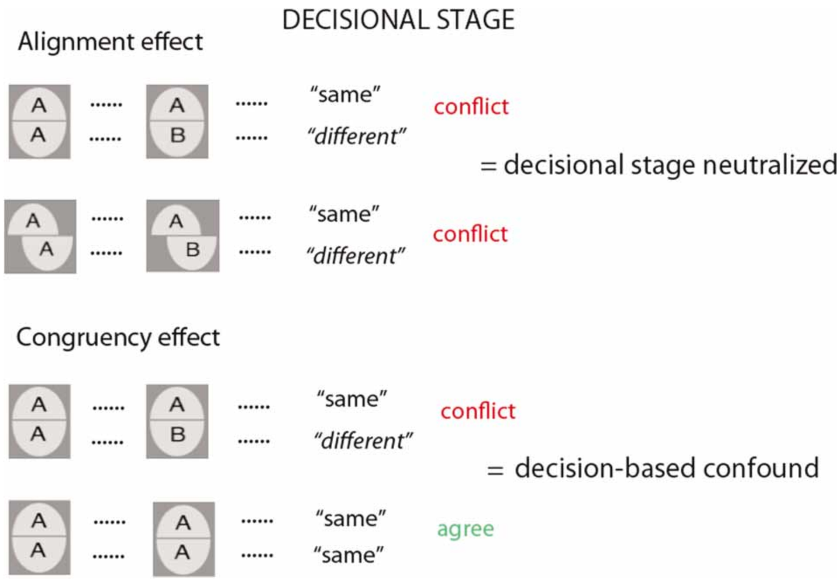

Figure 36. In the standard matching composite paradigm, which measures an alignment effect, the irrelevant part (the bottom face half) is associated with an expected behavioural response ("different") that enters in conflict with the expected behavioural response for the target in both conditions compared (aligned and misaligned faces). Hence, there is no built-in response conflict confound in this paradigm. However, in the congruency paradigm (shown here only for "same" trials, for simplicity), the irrelevant part (the bottom face half) is associated with an expected behavioural response ("different") that either enters in conflict or agrees with the expected behavioural response for the target. Hence, even if the two parts are processed completely in parallel, an effect of congruency can arise purely at the level of the output in this latter paradigm. Because of this response conflict confound inserted in the paradigm, decisional factors are very likely to contribute to congruency effects obtained in this paradigm, and the paradigm is heavily undetermined. This is an issue that plagues all of Gauthier and colleague's studies, as well as a few other studies on faces (see main text), but not the studies in which the context is behaviourally neutral, as in the Garner paradigm with faces (e.g., Amishav & Kimchi, 2010). To view this figure in colour, please see the online issue of the Journal.

However, what the field of face processing is particularly interested in is perceptual integration of face parts, namely holistic perception, rather than a general congruency/interference effect between face parts associated with additive/conflicting motor responses.

Missing a misaligned condition. Fortunately, most of the studies that measure an effect of context with face parts associated with competing behavioural outputs include a control condition such as inversion (Anaki et al., 2011; Farah et al., 1998; Goffaux, 2009, 2012). In these studies, if the congruency effect is compared for upright and inverted faces, the response conflict confound is neutralized (even though it would be better to avoid having such a built-in confound in the first place). In the studies of GRC with the congruency/interference composite face paradigm, such a control condition, inversion, or misalignment is therefore absolutely necessary. However, the authors have been very critical of the misalignment manipulation in many publications, claiming that "the congruency effect provides a single measure of holistic processing without necessitating a misalignment manipulation to measure it" (Cheung et al., 2008, p. 1328). That is, misaligned trials are not systematically included in this paradigm, so that congruency effects alone are interpreted as evidence of holistic processing (e.g., Bukach et al., 2006; Richler, Gauthier, et al., 2008, Exp. 1; Richler, Mack, et al., 2009; see also Curby, Goldstein, & Blacker, 2013). Moreover, even when misaligned trials are included, the main effect of congruency, independently of misalignment, is considered as reflecting "holistic processing". This way of proceeding concerns all the studies of Gauthier and colleagues on composite stimuli, whether these are faces or nonface objects.

Contrary to these authors, I argue that this response (in)congruency between the target and its context, coupled with a lack of a control (misaligned) condition, prevents making valid claims about holistic face processing, as defined in terms of perceptual integration. Indeed, if a control condition with misaligned (or inverted) trials is not included, how do we know that the performance decrease in incongruent trials is not due to the mere presence of *any kind* of physical difference that is present between the two visual displays to compare? One could potentially observe a decrease of performance in matching two identical top halves of faces if, below each half face presented in succession, there were a picture of a goat and then a picture of a rabbit (incongruent trials), rather than the same two pictures of a goat (congruent trials) (Figure 37). In these circumstances, indeed, having to discriminate two different half faces might also be influenced by whether the two animals are identical (incongruent trials) or different (congruent trials). A lower performance in incongruent than congruent trials would indicate that the animal picture interferes at some stage of processing with the target face half. However, one would never consider that such an effect reflects "holistic processing" in terms of "perceptual integration" of the top face and the animal. Rather, one would conclude that the presence of an animal interfered with the top face, and would then discuss the possible functional locus of this interference (likely to be attentional, at the level of the decision/motor response, or else because of competing independent perceptual representations).

This example shows that holistic processing cannot be assessed by comparing the processing of a target when there is "something else" (the context) in the display, which is associated with an (in)congruent behavioural response with respect the target. Indeed, an effect of congruency could emerge due to the context being processed independently from the target

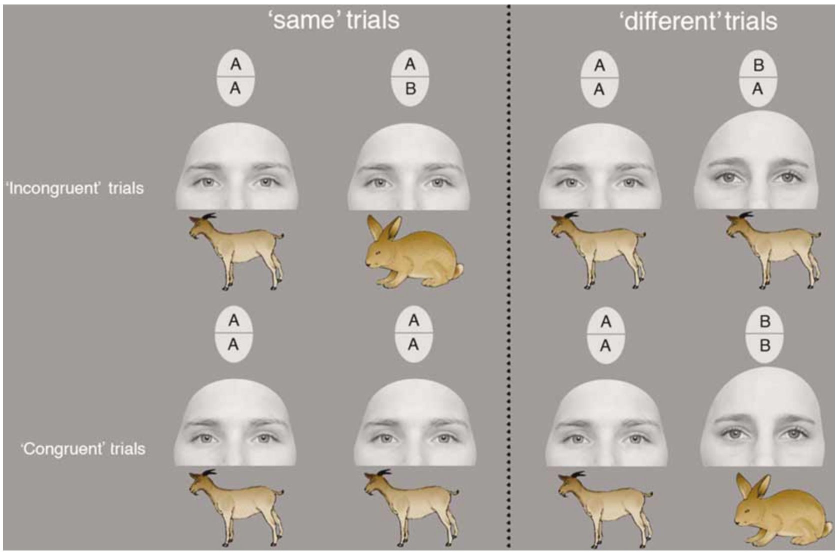

Figure 37. The congruency design revisited. The bottom halves of faces have been replaced by a picture of an animal, a goat or a rabbit, which could be congruent between trials, or incongruent. Judging whether the two faces' halves presented in succession are the same or different is likely to be influenced by the congruency of the two animals, especially since this "animal" context is, by itself, associated with a behavioural motor response. However, one would never conclude from such a reduced performance for incongruent as compared to congruent trials that the half face and the animal are integrated into a holistic representation. To view this figure in colour, please see the online issue of the Journal.

until the response stage. At the very least, one needs a control condition in which the two displays differ by a context that does not affect performance. This is exactly the reason for which misaligned trials are included in the standard composite design. Such a control condition is *necessary* to isolate holistic face processing in a congruency/interference face paradigm—it is not just a luxury add-on. For this reason, unless one prejudges the issue by arbitrarily deciding that holistic face processing is a general form of conflict between independent representations at any stage of processing including the motor response, conclusions about holistic processing that are made from composite face studies that do not include misaligned or inverted trials (e.g., Bukach et al., 2006; Richler, Gauthier, et al., 2008, Exp. 1; Richler, Mack, et al., 2009), or are based on congruency effects alone (virtually all of Gauthier and colleagues' studies with faces and objects), cannot be trusted.

Interference without integration: Two examples. The importance of misaligned trials can be illustrated by the study of a well-known case of acquired prosopagnosia. Bukach et al. (2006) tested the patient LR in the congruency paradigm of GRC without misaligned trials. The patient had a significant effect of *congruency* (congruent > incongruent), which was in the same range as the controls (Bukach et al. 2006, Exp. 2, Fig. 3). The authors concluded, and even stated in the abstract, that LR showed "normal holistic processing of faces". Note that this conclusion is at odds with a large number of studies showing that patients with acquired prosopagnosia present with impairments in holistic face processing (see earlier). Recently, we tested the very same patient LR and showed that he has impaired holistic processing of individual faces as assessed by the inversion effect, whole–part advantage (both weaker than normal controls), and even gaze contingency (Busigny et al., 2012). These observations are incompatible with Bukach et al.'s (2006) conclusions based on their congruency/interference paradigm. To clarify this, we tested LR with composite faces and also found a sort of congruency effect (i.e., the paradigm illustrated on Figure 24, and used in Busigny et al., 2010): In aligned "same" trials he tends to respond "different" more often when the bottom halves differ than when they are the same. However, contrary to controls, the patient also presents with such an effect for misaligned trials (Figure 38). Thus, LR appears to take into account the bottom half of the face to answer, even when it is misaligned with the top half (presumably because he has the tendency to overuse the mouth, as observed in other prosopagnosic patients; see Caldara et al., 2005, and Figure 30). Therefore, if the effect observed for misaligned trials is subtracted out from the effect observed for aligned trials, the patient's magnitude of the composite effect is below the range of the normal controls, an observation that is in line with the findings made with other paradigms, and supports an impairment in holistic face processing.

Another example comes from a study of Gauthier, Klaiman, and Schultz (2009), in which the authors tested individuals having an Autism spectrum disorder (ASD). With aligned stimuli, there was a congruency effect in individuals with ASD, like normal controls. However, whereas this effect disappeared for controls with misaligned faces, individuals with ASD showed the same congruency effect, regardless of alignment. Hence, if misaligned faces had not been used in that study, the authors' conclusion would have been that individuals with ASD present with normal "holistic processing" as measured by the congruency/interference paradigm. Therefore, this study directly contradicts these authors' own claims that misaligned trials are not necessary to draw conclusions with regard to the intactness of holistic processing.

Such a pattern of results illustrates why a misaligned trials condition—or another control condition such as inversion—is *necessary* to assess holistic processing with composite faces. If misaligned trials are not included, or not considered in the interpretation, an effect of congruency alone may have nothing to do with holistic face processing, defined in terms of perceptual integration. In fairness, GRC also included misaligned trials in several of

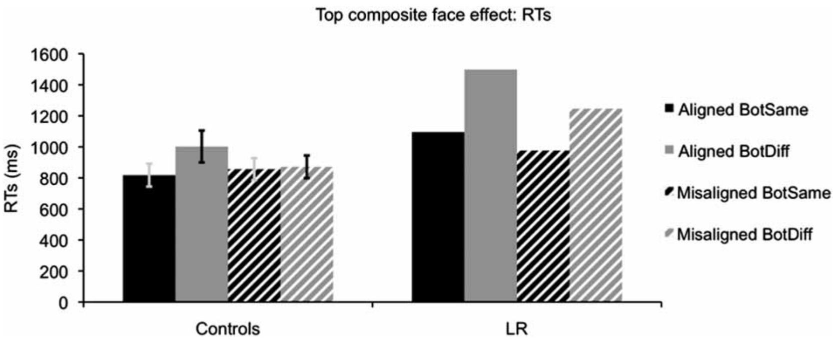

Figure 38. The patient LR's correct RTs data in the composite face paradigm ("same" trials only), compared to normal controls (Busigny et al., 2012). At first glance, LR presents with a composite face effect for top halves of faces, just like controls: He is slowed down on aligned trials when the bottom halves differ as compared to when it is identical. However, contrary to normal controls, he also shows the effect for misaligned faces, showing that such trials are necessary to avoid concluding erroneously that LR shows normal holistic face processing (Bukach et al., 2006).

their studies (Cheung et al., 2008; Gauthier et al., 2009; Richler, Cheung, & Gauthier, 2011a, 2011b; Richler, Mack, et al., 2011). However, even in these studies, the authors usually drew conclusions about holistic processing from a main effect of congruency. For instance, because misaligned trials gave rise to significant congruency effects at times of comparable magnitude to those observed for aligned trials, the authors claimed that misaligned faces were processed holistically (Richler, Tanaka, et al., 2008, Exp. 1). Moreover, despite the fact that congruency effects for nonface objects ("Greebles") were equally large for aligned and misaligned stimuli (Gauthier & Tarr, 2002; Gauthier et al., 1998), these authors concluded that these nonface objects were processed holistically. In reality, these studies reveal a general interference/congruency effect that can be found with pretty much any kind of visual display made of two congruent or incongruent elements. There is no way that such a congruency effect can be interpreted as reflecting perceptual integration between these elements.

A congruency effect on different trials reflects part-based processing. "Different" trials are not associated with a composite illusion (Figures 24 and 34), so that performance at matching different top halves trials should not differ when their aligned bottom halves are identical or different. Indeed, considering only the aligned trials in Gao et al.'s (2011) study, there is a congruency effect for "same" aligned trials (12.7% in accuracy), $t(23) = 38.81$, $p < .0001$; $t(23) = 4.39$, $p = .0002$. In contrast, there is only a small effect for "different" trials, in accuracy only (2%), $t(23) = 4.83$, $p = .04$; RTs, $t(23) = 0.57$, $p = .57$ (Figure 39).

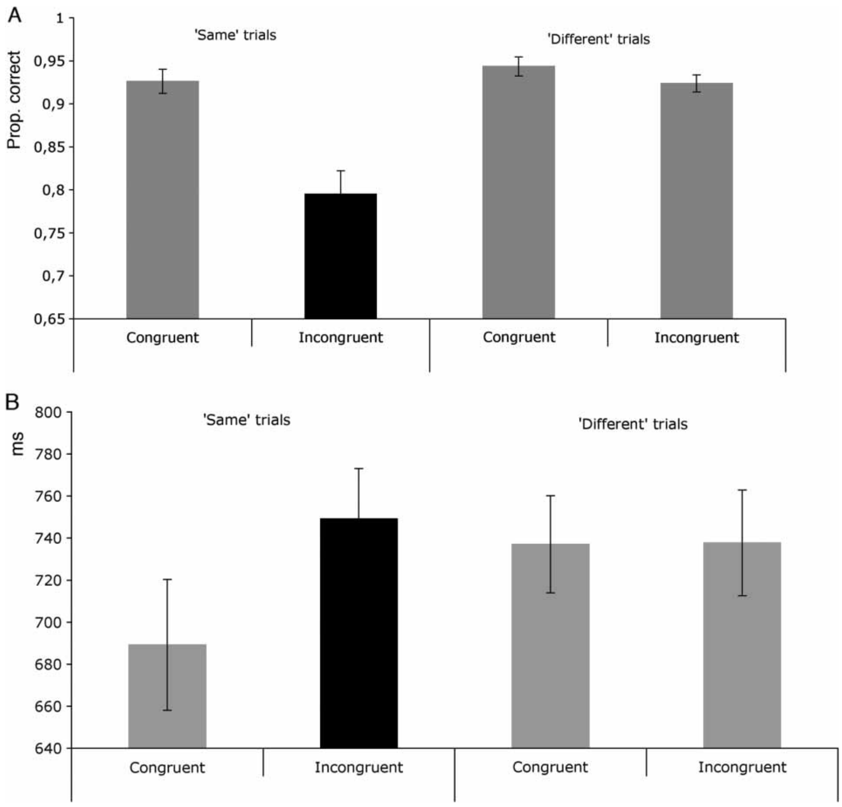

Figure 39. Data (24 participants, mean $\pm SD$) from the study of Gao et al. (2011). For incongruent trials, this is the same set of data as illustrated on Figure 25, but congruent trials have now been added. One can see that the only condition that differs from the others is the aligned "incongruent" condition for "same" trials, in terms of lower accuracy (A) and longer response times (B).

What does such a small effect on "different" trials, or any putative effect on "different" trials in such a congruency/interference design, really mean? This issue has already been discussed, but let me put it differently here. If one has to focus on the top but really uses the *whole* face to provide a response, the expected response is "same" for "same congruent trials", and "different" for "different congruent trials". For "same incongruent trials", the expected response is "different", leading to a drop in performance. However, critically, for "different incongruent trials", the expected response is also "different". That is, given that the *whole* face looks different in "different incongruent trials" (Figure 34, upper right corner; or Figure 24), if the face is processed holistically one should never expect a "same" response in such trials. If nevertheless, a "same" response is observed in a proportion of "different incongruent trials", it should never be interpreted as evidence for holistic face processing. Rather, this would indicate either that participants pressed the wrong button by mistake, or that for some reasons they were able to perform the task on the bottom half alone, i.e., *without using the whole face*.

I encourage the reader to pay attention to this issue because it is yet another important reason for which the results usually obtained with GRC's congruency/interference paradigm cannot be interpreted in terms of holistic processing. If one manipulates congruency on the "different" trials, these trials should not be included in the analysis. Alternatively, they could be used as a sort of control condition, reflecting part-based processes, in order to ensure that the participant did not simply base his/her decision on the sole irrelevant face half.

Unfortunately, GRC analyse their data with Signal Detection Theory (SDT, see earlier) and take the difference in d' between *all* the congruent and incongruent trials as an index of holistic face processing (e.g., Bukach et al., 2006; Curby et al., 2013; Cheung et al., 2008; Gauthier & Tarr, 2002; Richler, Cheung, & Gauthier, 2011a, 2011b; Richler, Gauthier, et al., 2008; Richler et al., 2009; Richler, Mack, et al., 2011; Richler, Tanaka, et al., 2008; Richler, Wong, & Gauthier, 2011; see also Goffaux, 2009, 2012; Meinhardt-Injac et al., 2010, 2011). Because mistakes made on "different" trials cannot, *in any way*, be associated with a decision based on the whole face (Figure 34), this d' congruency index can only reflect, at best, a diluted measure of holistic face processing. Worse, this index may be artificially inflated by the contribution of a putative effect of congruency on "different" trials, an effect that can only reflect a purely part-based process. Note that it is not SDT that is at fault here, and that this latter approach can be used with the standard composite face paradigm if one wishes to (see earlier). The problem arises when one manipulate congruency on "different" trials and include these part-based trials in the analysis to interpret them in terms of holistic processing. This is fundamentally incorrect.

Misinterpreting response bias. In their studies with the congruency/interference paradigm, GRC consider only the d' as being relevant for measuring holistic processing. The bias/criterion is considered as a "*response* bias" or a "*decision* bias", i.e., an effect of "cognitive/decisional" nature, whereas the d' measure would reflect "true discriminability". However, as explained previously, the bias/criterion could be as valid a measure as the d', and have a perceptual basis. Hence, dismissing an effect obtained in bias/criterion, as in GRC's studies, is missing the whole point. Note that since these authors' paradigm has a built-in response conflict confound, the bias as measured in SDT might indeed reflect at least partly a decisional/motor output process in their paradigm, contrary to the standard composite face paradigm. However, the contribution of perceptual and decisional/output factors cannot be disentangled by using such variables.

Summary. In summary, Gauthier, Richler, and colleagues developed their own version of the composite face paradigm, inspired by general congruency/interference paradigms such as the Stroop design, the Eriksen flanker task, or the Navon task. Although these latter paradigms are used to test for attentional interference processes and response conflicts, GRC used their congruency/interference composite face paradigm to study holistic face processing. They also sought to use that paradigm to replace the standard composite face paradigm, which does not have its roots in the attentional/decisional interference literature but emerges from a powerful visual illusion, and the phenomenology of face perception. Contrary to these authors, I argue that the two kinds of paradigm come from different sources and are so different that there is no point in wanting to replace one by the other one. In short, it is misleading to refer to the congruency design as being the "full design" and the standard composite face paradigm as a "partial design" (e.g., Gauthier & Bukach, 2007; Richler, Cheung, & Gauthier, 2011a, 2011b; Richler, Gauthier, et al., 2008; Richler, Mack, et al., 2011b). One should be called the standard composite face paradigm and the other one the congruency/interference composite face paradigm.

In this section, I have also argued that GRC's congruency/interference composite face paradigm has three major weaknesses that render it inadequate to make appropriate inferences about holistic face processing. First, the irrelevant part of the face ("context") is associated with a dual behavioural response that either conflicts or supports the dual behavioural response of the relevant part ("target"), so that congruency effects can arise due to response conflicts, in the absence of perceptual integration (as in the Eriksen task or when the Garner paradigm is used with conflictual responses in a same/different judgement task; Eriksen, 1995; Garner, 1988). Second, the paradigm typically lacks a control condition, such as misaligned or inverted faces, making the effects of congruency interpretable. Third, an effect of congruency on "different" trials does not reflect the processing of the stimulus as a whole but can only reflect part-based processes, so that including these trials to compute the effect is mistaken.

These fundamental problems concern all the studies of Gauthier, Richler, and colleagues with composite faces, but also other studies that used a similar paradigm with standard composite faces (Curby et al., 2003; DeGutis, Wilmer, Mercado, & Cohan, 2013; Gao et al., 2011; Xiao et al., 2012; Zhou, Cheng, Zhang, & Wong, 2012), or different face parts (Farah et al., 1998; Goffaux, 2009, 2012; Meinhardt-Injac et al., 2010, 2011). In the latter studies, control conditions such as misaligned faces (DeGutis et al., 2013; Gao et al., 2011; Xiao et al., 2012) or inversion (Goffaux, 2009, 2012) are included, so that these control conditions must be used to derive a measure that could be related to holistic face processing rather than to response conflicts. However, in order to draw proper conclusions about holistic face perception, the data of these studies should be reanalysed by using the "same" trials only.

Multiplying the chances to find "holistic processing". When GRC include misaligned trials in their studies (Cheung et al., 2008; Gauthier et al., 2009; Richler, Cheung, & Gauthier, 2011a, 2011b; Richler, Mack, et al., 2011), there are eight kinds of trials in total. This makes a paradigm that should not be called a "full" design, but rather an overextended design. Displaying the entire set of data of the study of Gao et al. (2011) illustrates this point (Figure 40): There is only one condition (aligned incongruent trials) that differs from all the others. That is, a single comparison to an appropriately matched control condition (misaligned incongruent trials) is sufficient to derive conclusions about holistic face processing (in the sense of perceptual integration). However, GRC do not only include both the "same" and the "different" trials: They consider an interaction between two factors (congruent (aligned > misaligned) > incongruent (misaligned > aligned)), but also a main congruency effect (congruent > incongruent), and an alignment effect (misaligned > aligned) as evidence for holistic processing. Basically, the paradigm provides three chances instead of one to observe an effect that is interpreted as evidence for "holistic processing". Therefore, it should not come as a surprise that evidence for "holistic processing" is found for pretty much any kind of stimulus in these authors' studies: Misaligned faces (Richler, Tanaka, et al., 2008), inverted faces (Richler, Mack, et al., 2011), nonface novel objects ("Greebles", Gauthier & Tarr, 2002; Gauthier et al., 1998; or "Ziggerins", Wong et al., 2009), nonface categories such as cars (Bukach et al., 2010; Gauthier, Curran, Curby, & Collins, 2003), English words (Wong et al., 2011), Chinese characters (Wong et al., 2012), or even musical notations (Wong & Gauthier, 2010).

GRC's overextended congruency design: Methodological confounds

At this point, I must make it absolutely clear that a congruency design that includes misaligned trials as control conditions is not methodologically incorrect. If one aims at measuring holistic processing, such a design is overextended, but it includes the appropriate conditions. Providing that the "different" trials in which congruency is manipulated are not included in the analysis, and that only the interaction between congruency and alignment on "same" trials is interpreted, one can make inferences about holistic face processing. For instance, the study of Gao et al. (2011), which has been used in this review to display data (Figures 25, 39, and 40), is methodologically sound in terms of data collection. What is problematic is the inclusion of "different" trials in the measure of holistic processing, so that the study's conclusion (i.e., that holistic face processing is primed by processing a Navon

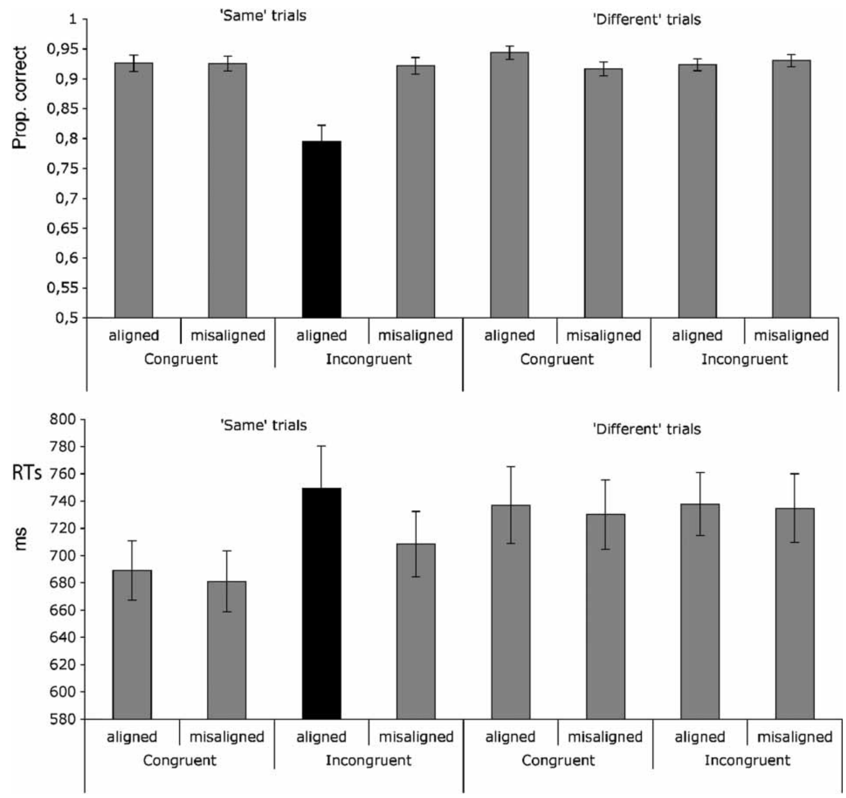

Figure 40. The full data set (24 participants) from the study of Gao et al. (2011). This is the same set of data as illustrated on Figures 25 and 39, but now it includes the congruent trial conditions. One can see that the only condition that clearly differs from the others is the aligned "incongruent" condition for "same" trials.

stimulus at the global level) might not be valid. Nonetheless, this can possibly be fixed because the data were acquired in a methodologically sound design. This is also true in the studies that used a congruency paradigm with facial parts and included the necessary control conditions (Anaki et al., 2011; de Gutis et al., 2013; Farah et al., 1998; Goffaux, 2009, 2012).

However, throughout all their studies, Gauthier, Richler, and colleagues have developed a particular version of the congruency/interference paradigm that has many additional methodological confounds, which in turn can lead to spurious effects and make unreliable conclusions. I would like to devote the present section to these methodological confounds, not only to show that the claims made by GRC about holistic processing are often invalid, but also to explain why despite being much less sensitive to holistic perception effects, their paradigm can lead to all sorts of spurious effects that can be found with any visual shape. More generally, these issues provide a very good illustration of the kind of problems that can arise if one reasons on this issue in terms of general attentional/response interference rather than on the phenomenology of face perception.

Stimulus confounds. The composite face paradigm is based on a visual illusion and is aimed at capturing this perceptual phenomenon. Given this, it is important to construct the composite face stimuli by trying to maximize the effect while controlling for potential pitfalls. The first thing to do is to ensure that the identical face halves to be matched—the top halves—really are physically identical. Given that faces vary a lot in height/width, creating a composite face made of the top of Face A and the bottom of Face B requires a careful adjustment of the width of the bottom half of B so that the top and the bottom halves form a continuous shape, that is, a "whole" face. Such composite faces should also be as realistic as possible so that the shape of the nose is relatively well preserved, as illustrated in the figures of this paper. However, when one reasons from an attentional perspective rather than considering a perceptual phenomenon, these issues may appear far less important. For instance, in all their studies, GRC do not adjust the bottom halves to the top halves of their face stimuli in a systematic (e.g., pairwise) way. Rather, a set of top parts is *randomly* combined with a set of bottom parts. Because of this random combination, the face stimuli used in these studies are generally inappropriate, or at least suboptimal, to capture perceptual integration effects (i.e., holistic perception). Let me illustrate this issue at three levels.

Misaligned aligned faces. The most salient problem comes from the study of Cheung et al. (2008): The faces were cut in two halves that were *randomly* combined. However, because the faces in that study differed in face width, such random combinations led to aligned composite faces in which the two halves did not fit at all (Figure 41, left side of Figure 2 in Cheung et al., 2008). In other words, aligned faces were somewhat spatially misaligned in that study, minimizing the contribution of perceptual integration of facial parts in any effect obtained.

The width of a circle. Perhaps to avoid this problem of a poor fit between top and bottom halves, GRC performed most of their subsequent studies (Richler, Cheung, & Gauthier, 2011b; Richler, Gauthier, et al., 2008; Richler, Mack, et al., 2009, 2011; Richler, Tanaka, et al., 2008, Exp. 3), with faces for which the width and height were normalized by applying the same oval shape to all faces. Using this procedure prevents at least the misfit of the top and bottom halves. However, there is a more fundamental, and in fact interesting, problem that arises when using such "circular" or "ovalized" faces: The composite face effect can be substantially reduced. Again, one has to turn to the composite face illusion to appreciate it: The strength of the

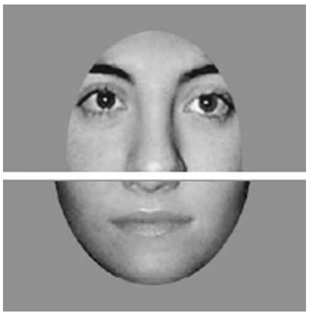

Figure 41. The kind of "aligned" stimuli used by Cheung et al. (2008); taken from Fig. 2 in that paper, original stimuli correctly aligned by Goffaux & Rossion, 2006). The bottom half is wider than the top half. This unfortunate misalignment in "aligned" trials is a consequence of random pairing of the top and bottom halves of a large set of different faces.

visual illusion is reduced when one normalizes the shape of the face in width and height by using the same oval shape for all faces (Figure 42).

This observation is not that surprising because the composite face effect, as measured in the standard way, depends mainly on shape-based information (Jiang et al., 2011). Among the variations of shape that are important, it is very likely that large-scale variations, such as the overall shape (i.e., the contour) or the height of the bottom half, are critical.[7] As an illustration, one can present the same top half associated with the exact bottom half that is simply stretched vertically (Figure 43). Because our face processing system is highly sensitive to the aspect ratio of individual faces (e.g., Barton et al., 2003; Haig, 1984; Lee & Freire, 1999), the two top halves are perceived as different (a composite face illusion). For instance, in this particular case, stretching the bottom half causes the visual impression that the eyes are closer to each other. Another good example, related to the composite face illusion, comes from the head size illusion (Morikawa et al., 2012; see Figure 43B), in which the lower part of the head influences the judgement of size of the upper head. These examples show that, if possible, the face stimuli used in a composite face paradigm should *not* be normalized to "eliminate the cues derived from the shape of the head or chin" (e.g., Richler, Mack, et al., 2009, p. 2856; and other studies). Otherwise, one minimizes again the contribution of perceptual integration of facial parts in the effects obtained. *Lumping together the top and bottom face halves trials does not help integration.* With unfamiliar faces at least, the composite face illusion is a phenomenon that is only clearly observed on the (identical) top halves, not

[7] This important role of the contour is also likely to be the reason why even when the internal "configural cues" of the top half are modified to make the face grotesque in an unchanged contour, there is still a composite face effect (de Heering, Wallis, & Maurer, 2012).

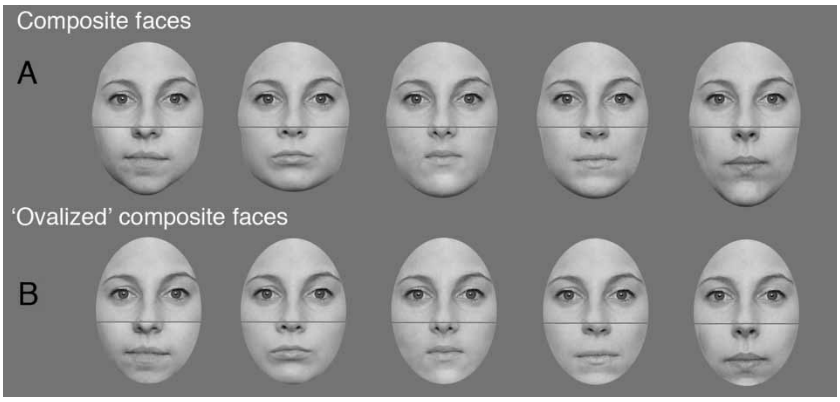

Figure 42. (A) Normal composite faces. (B) The exact same stimuli, which have been normalized, or "ovalized". Although the visual illusion (i.e., perceiving the top halves as being different) is present in both cases, it is certainly more compelling when the outline shape of the face is preserved, as in A. Therefore, in order to capture the perceptual phenomenon corresponding to the visual illusion, one does not want to normalize (i.e., eliminate) the shape of all faces in the experiment. The reason is that the composite face effect/illusion is primarily driven by shape variations, and variations of the height of the bottom halves of faces.

on the bottom halves (Figures 1 vs. 29, see earlier). The top half is more diagnostic of facial identity, and is naturally fixated by human observers. Hence, in almost all studies of the standard composite face paradigm, participants attend to only the top half. In the rare cases when the composite face effect is measured on the bottom half, the measure obtained on the top and bottom halves is not mixed up (e.g., Nishimura et al., 2008; Ramon et al., 2010; Young et al., 1987). However, probably because GRC do not reason in terms of a perceptual phenomenon, these authors manipulate congruency on both the top and bottom halves, and in the vast majority of their studies they lump together the "top" and "bottom" trials in a single measure. Given the asymmetry between the top and bottom halves of faces, this procedure can only lead, again, to a much weaker contribution of (holistic face) perceptual factors to the effects that they obtain.

Change of format confound. In standard composite face studies, the encoding and the test stimulus are *both* presented in the same format.[8] That is, both are aligned, or *both* are misaligned. However, in GRC's studies (for exceptions see Exp. 1 of Richler, Tanaka, et al., 2008, in which this factor

[8] There are a few exceptions to this rule, including one of our studies (Hugenberg & Corneille, 2009; Michel, Rossion, et al., 2006), and the first experiment of de Heering et al. (2007), as discussed later. However, unlike what is done in GRC's studies, the change of format for the misaligned condition only was never associated with a shift of position for that condition only (i.e., no attentional confound).

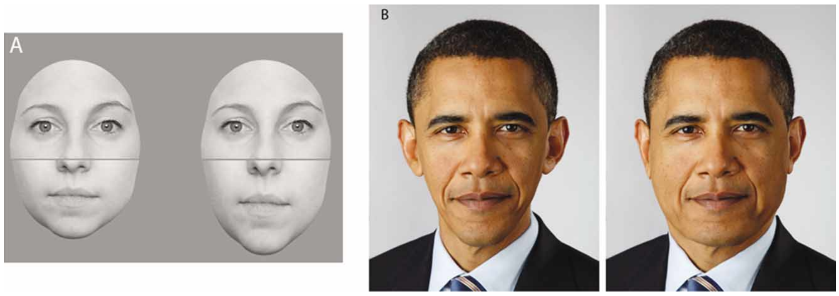

Figure 43. (A) In this example, the exact same top and bottom halves are presented. However, the bottom half is elongated in the example on the right side, creating a composite face illusion (the erroneous perception of the top halves as being different). Because our face processing system is highly sensitive to aspect ratio of individual faces, the eyes appear closer to each other, and the face or a smaller width, on the right stimulus. (B) The head size illusion (Morikawa, Okumura, & Matsushita, 2012), in which upper heads (above the eyes) of the same size are seen as different because the faces differ in width of cheeks, jaws, and necks. Fatter lower faces cause the head size to be 4% overestimated, and thinner lower faces cause the head size to be 3% underestimated. The illusion is dramatically reduced if the face is presented upside-down (Morikawa et al., 2012). This is the official president's photograph of the White House, which has been used in many press reports (Photo: The White House). This image is a work of an employee of the Executive Office of the President of the United States, taken or made as part of that person's official duties. As a work of the U.S. federal government, the image is in the public domain. To view this figure in colour, please see the online issue of the Journal.

was manipulated; and Cheung et al., 2008), there is no change of format in aligned trials but a systematic change of format in misaligned trials (Figure 44). Thus, the change of format is another factor that differs between the conditions (i.e., a methodological confound).

This confound can have two unfortunate consequences. First, because there is also an illusion of a perceived difference in the target part for "same" misaligned trials (see Figure 1 of Michel, Rossion, et al., 2006), the contribution of holistic face perception is reduced when computing the difference between aligned and misaligned trials. Second, this confound could even lead to incorrect conclusions. In the first experiment of the study of de Heering et al. (2007) with young children (4 to 6 years old), an aligned-to-misaligned stimulus presentation was also used in the control condition. Younger children (4–5 years old) tended to respond "different" to trials that had a change of format, leading to an unusually high rate of mistakes in that aligned-to-misaligned control condition, and consequently to the absence of a composite face effect. In contrast, adults or older children (6 years old) were not misled by the change of format. The investigators could have concluded that the composite effect emerges at 6 years of age. However, a second experiment with study and test faces presented without format change (both aligned vs. both misaligned) showed a large composite effect in 4-year-old children (de Heering et al., 2007). This example shows that

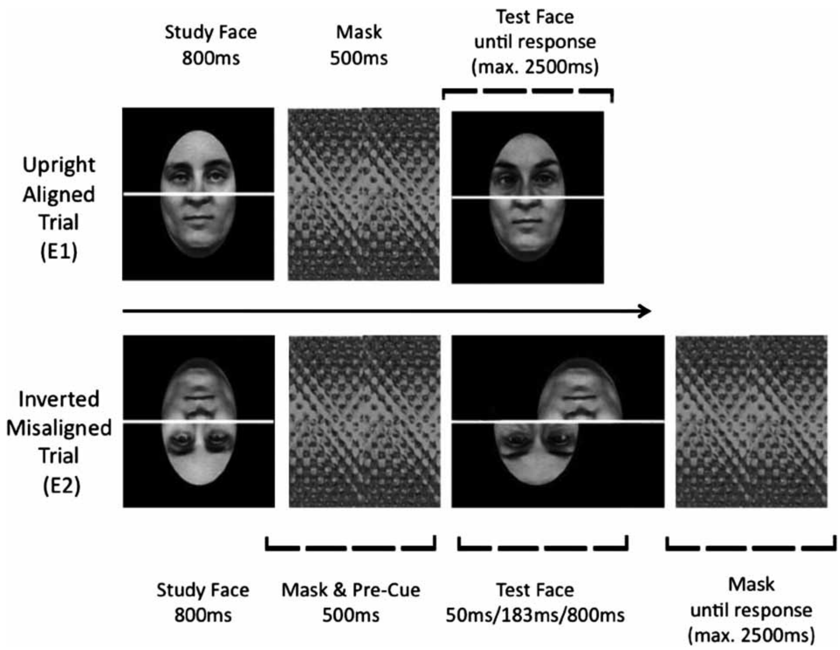

Figure 44. Examples of GRC's congruency paradigm (taken from Richler, Mack, et al., 2011), illustrating numerous problematic aspects of this paradigm. First, the comparison is made between an aligned-to-aligned and an aligned-to-misaligned condition, so there is a change of format between the two faces to be matched only in the misaligned condition (i.e., a methodological confound). Second, when the test face is aligned (top), the halves are not shifted laterally in position. However, when the face halves are misaligned, the parts are shifted laterally with respect to the first aligned face. Hence, the misaligned condition requires a shift of attention, and probably of eye gaze fixations, whereas the aligned condition does not (see also Figure 45). Third, in the upper display, the observer has 800 ms to encode the two half faces (contrary to encoding only the top half in the standard composite face paradigm) and does not know at this stage which one is going to be the target. Sometimes, the stimulus is presented for less than 200 ms, and there is only time to fixate one half face; sometimes the two halves can be fixated. Then, if the square bracket appears on the top of the test face, the participant should match/discriminate one of the two halves kept in memory with the top half face. If the bracket appears on the bottom, then the bottom half face should be used. Bottom (Exp. 2 of that study), the square bracket already appears in between the two faces, during the mask presentation. This peculiar procedure dramatically increases the complexity of the task, the working memory load, and the contribution of attentional factors to the performance.

changing format between the two faces to be matched, as done in all of GRC's studies when misaligned faces are included (see Figures 44 and 45), can lead to completely incorrect conclusions.

Spatial attention confounds

Lateral shifts of attention for misaligned trials. In all studies using the standard composite face paradigm, the top halves of the second (target) face always fall on the same fixation spot, whether the bottom halves are aligned or misaligned (Figure 2). Alternatively, if one changes the size or position between the encoding and test face, this manipulation is done to the same extent for aligned and misaligned faces, to ensure that one condition does not have an advantage over the other. In contrast, when GRC include misaligned trials in their congruency/interference paradigm, the two halves are shifted laterally with respect to the study face halves. This is not the case for aligned trials (Figures 44 and 45), another methodological confound. Because of that, eye movements and shifts of attention are likely to be artificially increased in misaligned compared to aligned trials, yet another—important—methodological confound. This point is illustrated on Figure 45 for one of GRC's studies with nonface objects, showing that a difference between aligned and misaligned trials could be due to a substantial spatial attentional confound.

Switching attention between top and bottom. In the standard composite face paradigm, the observer focuses on the top half of each of the two faces presented sequentially. The instruction given to the participant of such an experiment is clear: "Please focus on the top half of the face (above the white line separating the two halves) and decide if the top half is the same for the encoding and the test face that are presented in succession. Ignore the bottom half." There is evidence that participants respect this instruction, keeping gaze fixation on the top half (de Heering et al., 2008).

However, with the exception of one study (Cheung et al., 2008), GRC also introduced another significant modification in their congruency/interference paradigm. That is, participants have to consider *both* halves of the study face. *Then*, after it disappears, there is a cue indicating which of the two halves of the test face should be considered to make a decision (Figures 43 and 44). The cue is usually a square bracket surrounding the test face half to match (e.g., Figure 2 of Gauthier et al., 2009; Richler, Mack, et al., 2009; see also Figures 44 and 45). Sometimes the cue appears *at the same time* as the test face (e.g., Richler, Gauthier, et al., 2008; Richler, Mack, et al., 2011, Exp. 1; see also Curby et al., 2013), and sometimes it appears in the interval *in between* the two faces (e.g., Gauthier et al., 2009; Richler, Cheung, & Gauthier, 2011a, 2011b; Richler, Mack, et al., 2009, 2011, Exp. 2; Richler, Tanaka, et al., 2008, Exps. 1 and 2).[9]

Because of this manipulation, GRC's paradigm may lead to spurious congruency effects (i.e., better performance for congruent than incongruent trials only because of spatial attentional confounds). For instance, a participant who encoded both the top and bottom face halves will fixate first on the top half of the target face (e.g., Hsiao & Cottrell, 2008; Orban de Xivry et al., 2008). If the simultaneously presented cue indicates to use the

[9] Note that the cue is sometimes even presented after the test face, as in Richler, Tanaka, Brown, & Gauthier, (2008), Exp. 3).

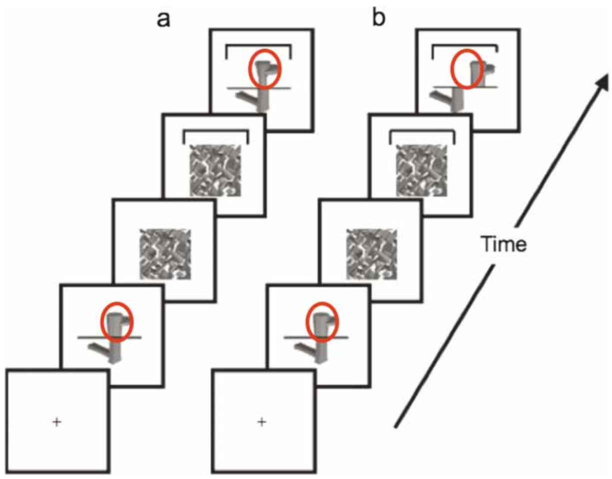

Figure 45. Examples of an aligned and a misaligned trial used by Wong et al. (2009) to measure the congruency/interference effect on nonface novel objects (Fig. 2 from that paper, red circles added by the author). Contrary to studies using the standard composite face paradigm, the authors use an aligned-to-misaligned trial in the "misaligned" condition (b). Critically, in this latter condition, the participant's fixation and attention has to shift to the left or right because the target part does not fall in the centre. It leads to an important spatial attention confound, so that some of the differences between aligned and misaligned trials in this congruency/interference paradigm could be due to attentional factors that have nothing to do with holistic processing. To view this figure in colour, please see the online issue of the Journal.

bottom half on this trial, the participant has to consider a nonfixated bottom half to make the decision. In this situation, when the test face appears, the participant ends up trying to ignore information from a part that he/she spontaneously fixated first, the top part. Consequently, the participant will certainly make more mistakes when there is an incongruent part *anywhere* in the second visual display presented (i.e., incongruent trials) than when everything in this display is congruent. Indeed, the first fixated part is incongruent only in incongruent trials.

Here is another problematic situation in GRC's paradigm that stems from the ambiguity as to which face half to encode. Sometimes presentation time may be limited at encoding so that there is no time to fixate the two halves of the study face (e.g., Hole, 1994: 80 ms). In the standard composite face paradigm the participant knows in advance which of the two halves (usually the top half) has to be encoded, and can therefore fixate gaze accordingly. Presentation duration can be the time needed for the target half face to be processed. In contrast, in GRC's congruency paradigm, if presentation time is short (<150–200 ms), he/she can only fixate one half. If presentation time is longer, he/she will probably alternate between fixating one of the two halves of the study face (increasing eye movements). Thus, in this congruency/interference paradigm, you cannot fairly compare the effects observed when presentation duration is variable at encoding, especially when you compare conditions allowing only one fixation (e.g., 17 ms to 183 ms) to conditions allowing several fixations (>183 ms until 800 ms) (Richler, Mack, et al., 2009).

Finally, as already explained, the results in GRC's paradigm may be highly dependent on one's fixation location preference for faces. Typical observers usually fixate first and primarily the top half of the face, but there is evidence that this is not the case for patients with acquired prosopagnosia, who preferentially fixate the mouth (Orban de Xivry et al., 2008; Van Belle et al., 2011). Therefore, any difference between normal observers and such patients in GRC's paradigm, especially when misaligned faces are not used, is enigmatic, and could be due to a different preferential spatial attention or fixation location. A similar reasoning may also be applied to populations with autism spectrum disorder, who present with a different pattern of gaze fixations on faces than typical observers (Klin, Jones, Schultz, Volkmar, & Cohen, 2002).

A too complex paradigm. Compared to the standard composite face paradigm, GRC's paradigm is complex, in terms of instructions, but also in the number of conditions used. In principle, if one includes misaligned trials, GRC's paradigm only has four experimental conditions (Congruency × Alignment). However, this is true if the top and bottom halves trials are grouped in the analysis, and the "same" and "different" trials are combined to provide a *d′* index. In reality, there are 16 different kinds of trials in GRC's congruency/interference paradigm, which are sometimes considered as 16 different conditions in the analysis by the authors (Wong et al., 2009). This number increases to 32 if inverted faces are included (Richler, Mack, et al., 2011). This is not parsimonious at all, and violates a general principle in research that one should include only the conditions in the paradigm that allow the testing of specific hypotheses rather than manipulate all possible variables and then expect some regularities ("laws") to emerge.

Such a high number of conditions is particularly problematic when one needs to assess holistic face processing in (very) young children (e.g., Carey & Diamond, 1994; de Heering et al., 2007; Macchi Cassia et al., 2009; Mondloch et al., 2007; Susilo et al., 2009), or in brain-damaged patients (e.g., Busigny et al., 2010; Ramon et al., 2010). These populations usually cannot be tested in long sessions, making it important to limit the number of conditions used to the most important ones (e.g., comparing only two conditions in the studies of de Heering et al., 2007, and Macchi Cassia et al., 2009, with the standard composite face paradigm). In these situations, a congruency/interference design with 16 types of trials is not realistic.

Moreover, such populations are usually slower than a typical adult population in terms of processing speed or response, and may show attentional and working memory limitations. In the standard composite face paradigm, when all participants of a given study know in advance that they have to focus and fixate a single element of a display (the top face half), the comparison between populations with different processing speeds and attentional/working memory capacities should still be reasonably fair. However, in an overextended congruency/interference paradigm, when participants are required to switch attention between two different face halves at encoding, atypical populations of participants might be particularly slow and thus fail to attend both halves of the face. Finally, when one has to take into account a changing cue in addition to the two face halves in each trial, the paradigm is not only heavily loaded in terms of attention and working memory, but the task must be very difficult to understand for a participant (just consider Figure 2 in Gauthier et al., 2009; or Figures 44 and 45 here). Besides the difficulty of ensuring that participants from atypical populations understand and perform such a complex task properly, differences between populations in the overextended congruency paradigm may thus emerge for reasons that have nothing to do with holistic face processing (general speed of processing, attentional maturation or defects, increased difficulties in response selection, etc.).

Summary. To summarize this section so far, Gauthier, Richler, and colleagues have developed an experimental paradigm with composite faces that is not tailored to capture a perceptual phenomenon. Rather, this paradigm measures the performance at processing a target and its context when they are associated with conflicting versus supporting behavioural outputs. This paradigm has important stimulus and spatial attentional confounds, ignores the asymmetry between the perceptual representation of the top and bottom halves of faces, and includes "different" trials reflecting part-based judgements in the critical measure. These manipulations undoubtedly reduce the contribution of a perceptual integration factor ("holistic perception") to the effects obtained in their paradigm. When misaligned trials were included, the authors also introduced novel methodological confounds: A change of format requiring lateral spatial shifts of attention for misaligned but not for aligned trials, which can lead to spurious effects of spatial attention in their paradigm. Because participants do not know which face half should be attended/fixated, this paradigm also increases working memory load, attentional resources, and eye movements, while making it absolutely inappropriate to manipulate the presentation time of the stimuli. For all these reasons, and also because the authors typically interpret three kinds of effects (misalignment, congruency, or the interaction between the two factors) instead of one, it is not surprising that GRC's particular congruency paradigm, albeit being much less sensitive to measure holistic perception, can lead to all sorts of differences between conditions that are interpreted by the authors as "evidence for holistic processing". For all the reasons detailed in this last section, I argue that these differences are not interpretable, and have probably only very little to do with holistic perception in the sense of perceptual integration.

GRC's criticisms of the standard composite paradigm: A short rebuttal. As mentioned earlier, Gauthier, Richler, and colleagues have been extremely critical of the standard composite face paradigm throughout the past few years. Most of these criticisms have been rebuffed (McKone & Robbins, 2011; McKone & Robbins, 2007). A full response to these criticisms is also provided indirectly in Parts 1 and 2 of the paper, so I will directly address these criticisms only briefly here.

GRC usually mention three "problems" with the standard measure of the composite face effect, and a fourth one more recently. The first one is that the composite face paradigm requires including misaligned trials so that it "roots the operational definition of holistic processing in one specific transformation—misalignment", something that is "empirically and theoretically problematic". Indeed, these authors consider that "misalignment is just one specific image transformation" and that "misalignment has no special experimental or theoretical status" (Cheung et al., 2008, p. 1328). For all the reasons developed already in this paper (see the Why Misalignment? section, in particular), I argue that this criticism of misalignment is incorrect.

According to GRC, the second "problem" of the standard composite face matching paradigm is the impossibility of examining false alarms and thus isolating "true discriminability" (d') from the bias/criterion, using signal detection theory. They generally reject the bias/criterion as being irrelevant for their measure of holistic processing, and consider differences in bias between conditions as being problematic in (their version of) the "partial" design (e.g., Richler, Cheung, & Gauthier, 2011b; Richler, Mack, et al., 2011). However, I have explained previously that although SDT *can* be applied to the standard composite face paradigm, this analysis has some limitations, and its outcome should not be misinterpreted. In particular, the bias/criterion of SDT can—and is likely to—have a perceptual basis in the standard composite face paradigm and is a highly relevant variable.

A third "problem" of the standard composite face matching paradigm is claimed to be that the incongruent trials are always associated with a correct "same" response while the congruent trials are always associated with a "different" response (Cheung et al., 2008). That is, response and congruency cannot be separated and that would be a "confound". GRC argue that this is a problem because "participants are more likely to respond 'different' on incongruent than congruent trials" and "this response bias could interact with other factors such as misalignment" (Cheung et al., 2008, p. 1329). This is the reason why GRC claim that the standard composite face paradigm has "poor construct validity" (e.g., Richler, Cheung, & Gauthier, 2011a). However, because congruency is not manipulated in the standard composite face matching paradigm, this criticism is irrelevant: There are actually no congruent trials! Congruency is a concept that was introduced by GRC in their own studies but in the standard composite face matching paradigm, according to GRC's terminology, all the trials are "incongruent" and differ only in term of spatial alignment of parts (Figure 4). Furthermore, if one adds so-called "congruent" trials to control for general effects of alignment, as we did in a number of recent studies (e.g., Jiang et al., 2011; see Figures 22 and 23), these trials are also systematically associated with a "same" response, so that there is absolutely no possible confound between response and "congruency".

A last criticism raised by GRC is that the congruency/interference paradigm, but not the composite face paradigm, correlates with face recognition performance (Richler, Cheung, & Gauthier, 2011b). I have already discussed in Parts I and 2 of this review why the composite face effect should not necessarily be correlated with face recognition performance, and why such an absence of correlation would *not* mean that holistic processing is not important for face recognition or that, to use GRC's own words, "our efforts at understanding holistic face processing constitute wild goose chases" (Richler, Cheung, & Gauthier, 2011b, p. 464). I have also referred to two recent empirical studies that directly contradict these authors' claims (see Avidan et al., 2011; Wang et al., 2012). To be honest, I do not know why the measure in GRC's congruency/interference paradigm correlated with face recognition performance in their particular study (Richler, Cheung, & Gauthier, 2011b), an effect that was recently replicated with a weaker correlation by De Gutis et al. (2013) and not replicated at all in a study that used the exact same paradigm as Richler et al.'s study (Zhou et al., 2012). The outcome of this paradigm is so dependent on general factors such as working memory capacities, spatial/selective attention, and response selection that getting such a correlation in a particular study may not be very surprising: It is likely to be driven by such general factors thus, it is very difficult to see how such a correlation would help understanding holistic face processing, and why such a finding would "salvage the central role of holistic processing in face recognition" (Richler, Cheung, & Gauthier, 2011b). After all, across individuals, memory for cars correlates significantly, albeit weakly, with memory for faces (Dennett et al., 2011), and upright and inverted unfamiliar face matching correlate even more (Megreya & Burton, 2006). Does this mean that upright and inverted faces are perceived in the same way? Researchers in this field are not interested in the general processes that can drive such interindividual correlations, but rather in what specifically differs between upright face processing, namely holistic face perception.

Finally, note that the paradigm that GRC label the "partial design" in their own studies is not the standard composite face paradigm. It is their congruency/interference paradigm that the authors analyse by considering "same" trials only (e.g., Richler, Cheung, & Gauthier, 2011a; Richler, Mack, et al., 2011). This paradigm includes all the methodological confounds described previously. Hence, one should not be misled by these authors' claims about results obtained with their "partial design": These results are irrelevant for studies using the standard composite face paradigm.

Unfounded claims from using the overextended congruency design

Before concluding, I would like to address the question of the convergent validity of GRC's congruency paradigm. What has been found with this paradigm, and how do these findings and their interpretations stand with respect to the (holistic) face processing literature?

A decisional locus for holistic processing?. In their most cited paper (Richler, Gauthier, et al., 2008), GRC attempted to identify the functional locus of their congruency effect by means of a multidimensional generalization of signal detection theory called general recognition theory (GRT; Ashby & Townsend, 1986). They asked participants to perform a matching task in which attention is not focused on one half of the face (at any point in the paradigm), and to judge the same/different status of both halves on every trial. This task is called "the complete identification task" and the authors claimed that it is only by using such a task that one can isolate indexes in the behavioural measure that reflect perceptual effects ("violations of perceptual separability", "perceptual independence"; PS and PI, respectively) or decisional effects ("decisional separability"; DS). I cannot go into detail here on using the GRT approach to test the contribution of decisional versus perceptual factors on performance at face matching tasks, but the reader is referred to Richler, Gauthier, et al., 2008, and also to other studies that have adopted it (e.g., Wenger & Ingvalson, 2002, 2003; see also Cornes, Donnelly, Godwin, & Wenger, 2011, for its application to the "Thatcher effect"). Running that "complete identification task" task with aligned and misaligned faces, the authors found "limited violations of PI" and "inconsistent violations of PS", but "clear violations of DS". Since these observations were in line with previous studies using GRT in the context of the whole–part paradigm (Wenger & Ingvalson, 2002, 2003), these authors claimed that there was little support for a perceptual encoding locus in the task, and that "holistic effects in face processing are decisional (Richler, Gaulthier, et al., 2008, p. 341)". That is, they concluded that faces were not perceived holistically, but rather that holistic effects in face processing arise because of processes occurring at the decisional level (Richler, Gauthier, et al., 2008).

What the authors exactly means by decisional level is not very clear because they sometimes seem to distinguish completely the functional locus of a decision from the locus of an interference at the response level (Richler, Cheung, et al., 2009[6]). Notwithstanding the fact that the built-in response conflict confound is inherent to these authors' matching paradigm, it is indeed possible that additional decisional factors, whatever they might be, play a role in the effect that they report. This would be in line with the distinction that these authors make in all their studies between the interpretation of the d' and the response bias/criterion, with the second index considered—erroneously—by these authors as reflecting exclusively a bias of a *decisional* nature. Interestingly, in a more recent study (Mack, Richler, Gauthier, & Palmeri, 2011), the authors themselves acknowledged that the GRT framework (Ashby & Townsend, 1986) was unable to accurately characterize the perceptual versus decisional source of simulated known instances of violations of perceptual or decisional separability. Thus, they acknowledged that critical cases of violations of perceptual separability are often mischaracterized in this framework as violations of decisional separability. That is, the study of Mack et al. (2011), entitled "Indecision on Decisional Separability", dismissed entirely the conclusions reached by Richler, Gauthier, et al. (2008) that holistic processing has a decisional locus. In other studies, these authors also appeared to change their view about the functional loci of the congruency/interference effect as measured in their overextended design, switching between perceptual (Richler, Mack, et al., 2009), attentional (Richler, Tanaka, et al., 2008), and decisional loci to the point where it is impossible to know what their real position on this issue is.

In fact, given what has already been discussed, it is not surprising that the authors are quite inconsistent about the source of their effects. Indeed, in *their* congruency/interference paradigm with composite faces, there are probably many factors (perceptual, attentional, working memory, decisional/ response, etc.) that contribute to the behavioural effect. The problem arises when the authors attempt to determine which of these factors account for "holistic processing" in their paradigm. It is impossible to do this with such an undetermined paradigm and this whole research enterprise does not appear to have advanced our understanding of the functional locus of holistic face processing at all.

At this point, it is perhaps worth reminding the reader that the reasoning on holistic face processing and composite faces started with a compelling visual illusion (Figure 1A), showing that two top halves of a composite face *are perceived* as different if the lower halves differ. One can see it, even without making any decisions (just look at the figures, for instance Figure 1, and please do not press any key on your keyboard). Insert this visual illusion in the context of a matching task on the top halves, and you find that observers "fall into the trap" and they answer as if the two top halves are indeed different. This is a simple demonstration of a visual illusion that drives an incorrect behavioural response. As illustrated earlier, the Müller-Lyer illusion (Figure 27), could also be embedded in a same/different behavioural paradigm, in which observers would tend to incorrectly answer "different" for two arrows of equal length that are associated with diverging or converging heads (a "same" trial, giving rise to a "different" response, as in the composite face effect). Does it mean that the functional locus of the Müller-Lyer illusion could be decisional? To me, making such claims reflect a deep misconception about what can be inferred from psychophysical data alone. It is of the same order as claiming that because participants of a given study responded with their right finger, then the locus of the effect should be in their right finger. Or, that if participants were asked to respond verbally, then that the functional locus of the effect would be in their vocal cords. Again, with the composite face illusion, we have to deal with a visual illusion that appears to be created by an integration of the bottom face half with the top half. Since it is a visual illusion, it is reasonable to think that its locus must be in the visual system, somehow. As for many other visual illusions, it reflects a perceptual inference, or a construction (Gregory, 1997), showing that our internal models of the visual world influence what we see (what is usually referred to as "top-down" processes, but is essentially a characteristic of high-level vision). If one wants to identify the functional locus of the composite visual illusion, there are more direct ways than making inferences from behavioural studies alone, such as neurophysiological measures, with or without behavioural correlates (Figures 14 and 15).

Prosopagnosia. I have already discussed this issue in another context (se earlier), so I will be brief here. Bukach et al. (2006), using the overextended congruency paradigm, concluded that the prosopagnosic patient LR had preserved holistic face processing. This conclusion is not only at odds with studies performed on other cases of prosopagnosia (see earlier), but it is not supported by evidence collected on the very same patient, showing that he is clearly impaired at holistic processing of individual faces as assessed by the inversion effect, whole–part advantage (both weaker than normal controls), gaze contingency, and even the standard composite face paradigm (Busigny et al., 2012; see Figure 6). It is also worth adding that in a more recent study, Bukach et al. (2012) even acknowledged that the patient LR was impaired at holistic face processing, contradicting their own previous conclusions.

Exposure duration. In one of their studies, GRC found an effect of congruency for exposures as rapid as 50 ms, claiming to demonstrate that holistic processing of faces emerges for very briefly presented faces (Richler, Mack, et al., 2009). Although this may be true, as demonstrated originally by Hole (1994, Exp. 2) for simultaneous presentations of two faces for 80 ms, the Richler, Mack, et al. (2009) study has many other limitations preventing such conclusions. Most importantly, unlike Hole (1994), Richler, Mack, et al. did not use any misaligned faces or inverted faces as a control, which makes it impossible to interpret their effects. Also, although they asked their participants to encode both face parts, the two face parts could be fixated in turn only at presentation durations sufficient to permit a saccade. Therefore, the conditions with long durations of stimulus presentation cannot be fairly compared to the conditions with short durations (<200 ms). Finally, Richler, Mack, et al. did not consider RTs as a measure of holistic face processing, despite the fact that RT differences between their congruent and incongruent trials varied substantially across stimulus durations. Ignoring RTs is problematic because in his seminal study, Hole found that at long durations of presentation the effect was present only in accuracy rates, whereas at short durations the effect emerged in correct RTs and was no longer significant in accuracy rates.

''Holistic'' processing of inverted faces. GRC found that the magnitude of the effect measured in their paradigm did not differ between upright and inverted faces, concluding that inverted faces were processed as holistically as upright faces (Richler, Mack, et al., 2011). This conclusion goes against numerous studies that found either an absence or a massive reduction of the standard composite face effect with inversion (e.g., Carey & Diamond, 1994; Goffaux & Rossion, 2006; Mondloch & Maurer, 2008; Robbins & McKone, 2003; Rossion & Boremanse, 2008; Young et al., 1987) and is of course incompatible with the disappearance of the composite face illusion with inversion (Figure 6). It also goes against a whole tradition of research showing that inverted faces are not, or only weakly, processed holistically. These studies used various paradigms such as matching of faces varying in one of multiple dimensions and analyses of interactivity through multidimensional scaling (Sergent, 1984), the whole–part advantage (Tanaka & Farah, 1993), or gaze contingency (Van Belle, de Graef, Verfaillie, Rossion, & Lefèvre, 2010; see Figure 8). In fact, even studies that used congruency/interference face paradigms that did not include all of the methodological shortcomings reviewed earlier, have found either no effects of congruency for inverted faces or much weaker effects for inverted faces than upright faces (Anaki et al., 2011; Farah et al., 1998; Goffaux, 2009, 2012). Importantly, inverted faces are usually presented for reasonably long durations in all these studies (a few hundreds of milliseconds), so that

Richler, Mack, Palmeri, & Gauthier's (2011) claim that inverted faces would not be processed holistically only at short durations (Exp. 2 of that study) is clearly inconsistent with the literature.

Object processing and visual expertise. Finally, it is worth mentioning that GRC's congruency/interference paradigm was initially developed to study the processing of nonface objects (Gauthier & Tarr, 2002; Gauthier et al., 1998), and was applied only later to faces (Bukach et al., 2006). In their first study with composite stimuli (Gauthier et al., 1998), the authors used these novel multipart objects called "Greebles". They acknowledged that they were unable to find differences between aligned and misaligned Greebles (i.e., composite effects), whether they tested novices or experts (Gauthier et al., 1998, p. 2416). For this reason, a familiar recognition task with composite Greebles was designed. In the recognition task, participants had to recognize half of a Greeble X; the other half was either from another Greeble Y ("incongruent trials", or composite Greebles), or from the same Greeble X ("congruent trials", or original Greebles). Participants performed better with congruent than incongruent Greebles (a congruency effect). This is not surprising because participants had two parts to help them make the correct decision in congruent trials versus one in incongruent trials. Critically, there was no significant interaction with the *alignment* of the two Greeble halves, indicating that the participants—who were all "Greebles experts"—simply used the two halves in an additive way, without integrating them at all, to improve their performance. Despite the absence of an effect of *alignment*, or of any interaction between *congruency* and *alignment*, the authors concluded that it "obtained the composite effect with Greebles" (Gauthier et al., p. 2418). Thus, the authors considered from the outset that the effect of congruency reflects "holistic processing" rather than the effect of alignment, a clear demarcation from the standard composite face paradigm.

In a subsequent study, a delayed matching task was used, with congruent and incongruent Greebles, misaligned and aligned (Gauthier & Tarr, 2002). This is the first use of their typical congruency/interference composite paradigm discussed in Part 3 of this review. Participants were tested before, during, and after training (five sessions of testing). None of the effects (congruency, alignment, and their interaction) were significant. There was a nonsignificant trend for an interaction between session and congruency because there was an effect of congruency (composite vs. original) after training (experts) only. However, there was no significant effect of alignment, nor any interaction between alignment and congruency.

In the next study that used composite objects of expertise (cars, Gauthier, Curran, Curby, & Collins, 2003), alignment was no longer manipulated and participants had to match the bottom parts of cars, ignoring the top parts. The authors reported a main effect of expertise and a main effect of congruency, but no interaction between the two. More recently, a group of participants trained with another set of three-dimensional objects, the "Ziggerins", were tested in the congruency/interference paradigm (Wong et al., 2009). On sensitivity (d'), there was a main effect of congruency, which did not differ between experts and novices. On RTs, there was a nonsignificant trend of an interaction between expertise, congruency, and alignment, which was not due to a slower response for the aligned incongruent condition, as predicted, but to a faster response to aligned congruent condition than all other conditions, only for experts. However, this advantage for the congruent aligned condition was larger in novices than experts on sensitivity, pointing to a speed–accuracy tradeoff in the task (Figure 5 of Wong et al., 2009).

Overall, these studies reveal an effect of congruency with composite stimuli for different kinds of nonface objects, *regardless of visual expertise*. Only one study found a nonsignificant trend for a larger effect for experts than novices (Gauthier & Tarr, 2002; but see Hsiao & Cottrell, 2009, for the opposite effect). Most importantly, besides the fact that these studies are quite inconsistent with each other in terms of the variables manipulated (alignment or not, focus on top or bottom half or both, transformed car stimuli with flipped top cars in Gauthier, Curran, Curby, & Collins, 2003, etc.) and the dependent variables considered (accuracy, sensitivity, RTs, . . . without any efficiency measure computed to take tradeoffs into account), there was *never* a significant main effect of alignment for nonface objects of expertise in these studies, nor a significant interaction between alignment, congruency, and expertise.

In short, whereas the effect of alignment truly appears to be specific to faces, the effect of congruency is observed for pretty much everything, including English words (Wong et al., 2011), Chinese characters (Wong et al., 2012), or even musical notations (Wong & Gauthier, 2010). This is problematic because if one aims at demonstrating that faces are not special, in particular that faces do not call upon specific holistic processes, the proper approach is to test visual experts with nonface object categories by means of the very same paradigm used to obtain the strongest face-specific effects in novices. If the standard composite effect of alignment cannot be found with nonface objects of expertise (see also Robbins & McKone, 2007), either the visual expertise hypothesis has to be rejected, or the stimuli and the training regime have to be improved. Dismissing that standard paradigm to replace it by a paradigm that measures a general effect of congruency is not appropriate, and seems circular.

To summarize this section, Gauthier and colleagues have developed their own alternative version of the composite paradigm, measuring congruency/

interference between parts, after failing to disclose the standard composite effect of alignment with objects of expertise. The congruency/interference paradigm was then extended to faces in a large number of behavioural studies, the authors making claims about holistic face processing that are often based on overinterpretations and do not agree with the literature. Given the numerous methodological problems of this paradigm as described in the previous two sections, it is rather reassuring that its convergent validity is weak, and observations made with this paradigm should not impact the wider literature on (holistic) face perception.

GENERAL CONCLUSIONS

Following a selective review of the composite face paradigm in Part 1, I divided this paper into two further parts. Part 2 is about the standard composite face paradigm. Part 3 is about a congruency/interference composite face paradigm. Over the past few years, the bottom part has interfered a lot with the top part. This is rather unfortunate, yet it does not mean that the two parts should be integrated into a single theoretical framework. Rather, following this review, I hope that the standard composite face paradigm will truly come out on top and that the congruency/interference face paradigm will no longer just sit at the bottom, but simply disappear from the field altogether. Indeed, using this latter paradigm can only serve to create confusion in the minds of researchers inside and outside of this field. The standard composite face paradigm has been unfairly labelled by Gauthier, Richler, and colleagues as a "flawed" paradigm in many publications; I have demonstrated here that it is these authors' own congruency/interference face paradigm that is counterintuitive, does not test what it claims to be testing, is replete with methodological confounds, is undetermined, and overly general.

In trying to understand the reasons why the congruency/interference paradigm is inadequate to measure holistic face perception, we have seen that it was inspired by, and belongs to, a general class of congruency/interference paradigms used for decades in experimental psychology, such as the Stroop design, the Eriksen flanker task, or the Navon task. Although these latter paradigms are typically used to test for attentional interference processes and response conflicts, and they have proved of value in certain areas of face research (de Haan et al., 1987; Young et al., 1986), Gauthier and colleagues extended this approach into a congruency/interference paradigm with composite stimuli in order to make inferences about perceptual integration of parts. However, this latter approach is doomed, mainly but not only because the irrelevant part ("context") is associated with a dual behavioural response that either conflicts with or supports the dual behavioural response of the relevant part ("target"), so that the effects obtained could entirely due to a response conflict even in the absence of any perceptual integration between (face) parts. This issue is extremely interesting from my point of view (see also Garner, 1988), and, given the wide use of congruency/interference paradigms in experimental psychology and cognitive neuroscience, it has implications that go well beyond understanding of the nature of holistic face perception.

In the first part of the review, I tried to convince the reader that, thanks to Young and colleagues (1987), we have in our hands a rich composite face paradigm, directly inspired from a visual illusion. In the tradition of the phenomenological approach to visual perception, which traces back to the Gestaltists at the beginning of the twentieth century (Köhler, 1929; Wertheimer, 1912; see Wagemans, Feldman, et al., 2012), I argue that this paradigm is a fantastic tool to study face perception, even though it should be improved in stimulus design and systematic parametric manipulations. It should also be complemented by other behavioural approaches—for instance using gaze-contingent face stimulation (Van Belle, de Graef, Verfaillie, Busigny, & Rossion, 2010; Van Belle, de Graef, Verfaillie, Rossion, & Lefèvre, 2010).

One of the greatest challenges of visual perception research is to merge this phenomenological approach with a neurophysiological approach (Spillmann, 2009) initiated at the beginning of the second half of the twentieth century (Hubel & Wiesel, 1962; Jung, von Baumgarten, & von Baumgartner, 1952). Understanding how the human brain builds a unified face percept is one of the greatest challenges of visual neuroscience because, in the early stages of visual processing, a face is represented by neurons with small receptive fields. These neurons provide information about local elements of the face, and these elements need to be combined—for instance, by convergence of inputs to neurons in higher order areas with larger receptive fields, or/and by temporal synchronization of the activity of distributed populations of neurons in lower level areas (the "binding" problem in vision; Treisman, 1996). Obviously, this question is not specific to faces and concerns a general problem for understanding human vision (Spillmann, 1999; Spillmann & Ehrenstein, 1996). However, it is with faces that the challenge is perhaps the most difficult, for a number of reasons: Faces are made of multiple elements arranged over a continuous texture space; these elements and the face as a whole are dynamic and continuously changing; faces are highly familiar stimuli for which top-down processes and representations are constantly at play during their perception; and faces need to be perceived at a sufficiently fine-grained level of resolution to be distinguished from one another. In summary, the human face may be the quintessential whole, or Gestalt (Pomerantz & Kubovy, 1986). Given these reasons, it is not surprising that face perception is subserved in the human visual system by a widely distributed network of populations of neurons occupying much of the ventral part of the (right) occipitotemporal cortex, from the occipital pole to the temporal pole (e.g., Haxby, Hoffman, & Gobbini, 2000; Rossion, Hanseeuw, & Dricot, 2012; Sergent, Otha, & McDonald, 1992; Weiner & Grill-Spector, 2010), a factor that also increases the difficulty of understanding how a unified face percept is built by the human brain.

To clarify this issue, experimental psychologists and cognitive neuroscientists have at their disposal a formidable tool, in the form of the composite face illusion and its disruption by manipulations such as a slight spatial misalignment of its parts or inversion. The visual illusion shows that the elements of a face, which are processed initially by different populations of neurons in the human brain, are tightly linked perceptually. Therefore, it is not surprising that the composite face illusion has been adapted to a methodological paradigm aimed at measuring how a nonfixated face (bottom) part is perceptually integrated with a fixated (top) part. As I tried to illustrate throughout this (critical) review, the paradigm has been used with many types of stimulus transformation, in different populations (infants and children, patients with prosopagnosia, nonhuman primates, etc.), and also to record spatiotemporal correlates of a holistic face representation in the human brain with methods such as fMRI and scalp ERPs. Collectively, these studies provide some insights into our understanding of holistic face processing, even though the challenges remain significant. For instance, we still do not have any objective trace of a holistic face representation in the human brain, we do not know if parts are represented in face-specific cortical areas independently of whole face representations and we lack direct evidence that the whole face is different than the sum of its parts, in a Gestaltist sense. These issues are extremely difficult to resolve and will require much further collective work, using a holistic approach to face perception.

REFERENCES

Abbas, Z.-A., & Duchaine, B. (2008). The role of holistic processing in judgments of facial attractiveness. *Perception*, *37*, 1187–1196. doi:10.1068/p5984

Altmann, C. F., Bülthoff, H. H., & Kourtzi, Z. (2003). Perceptual organization of local elements into global shapes in the human visual cortex. *Current Biology*, *13*, 342–349. doi:10.1016/S0960-9822(03)00052-6

Amishav, R., & Kimchi, R. (2010). Perceptual integrality of componential and configural information in faces. *Psychonomic Bulletin and Review*, *17*, 743–748. doi:10.3758/PBR.17.5.743

Anaki, D., Nica, E. I., & Moscovitch, M. (2011). Automatic aspects in face perception: Evidence from mandatory processing of distractor facial components. *Experimental Psychology*, *58*, 4–18. doi:10.1027/1618-3169/a000061

Andrews, T. J., Davies-Thompson, J., Kingstone, A., & Young, A. W. (2010). Internal and external features of the face are represented holistically in face-selective regions of visual cortex. *Journal of Neuroscience*, *30*, 3544–3552. doi:10.1523/JNEUROSCI.4863-09.2010

Ashby, F. G., & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review*, *93*, 154–179. doi:10.1037/0033-295X.93.2.154

Avidan, G., Tanzer, M., & Behrmann, M. (2011). Impaired holistic processing in congenital prosopagnosia. *Neuropsychologia*, *49*, 2541–2552. doi:10.1016/j.neuropsychologia.2011.05.002

Bartlett, J. C., Searcy, J. H., & Abdi, H. (2003). What are the routes to face recognition? In M. Peterson & G. Rhodes (Eds.), *Perception of faces, objects, and scenes: Analytic and holistic processes* (pp. 21–52). Oxford: Oxford University Press.

Barton, J. J., Press, D. Z., Keenan, J. P., & O'Connor, M. (2002). Lesions of the fusiform face area impair perception of facial configuration in prosopagnosia. *Neurology*, *58*, 71–78. doi:10.1212/WNL.58.1.71

Barton, J. J., Zhao, J., & Keenan, J. P. (2003). Perception of global facial geometry in the inversion effect and prosopagnosia. *Neuropsychologia*, *41*, 1703–1711. doi:10.1016/S0028-3932(03)00115-5

Baudouin, J.-Y., & Humphreys, G. W. (2006). Configural information in gender categorisation. *Perception*, *35*, 531–540. doi:10.1068/p3403

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

Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience*, *8*, 551–565. doi:10.1162/jocn.1996.8.6.551

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, *94*, 115–147. doi:10.1037/0033-295X.94.2.115

Biederman, I., & Kalocsai, P. (1997). Neurocomputational bases of object and face recognition. *Philosophical Transactions of the Royal Society of London: Biological Sciences*, *352B*, 1203–1219. doi:10.1098/rstb.1997.0103

Bindemann, M., Scheepers, C., & Burton, A. M. (2009). Viewpoint and center of gravity affect eye movements to human faces. *Journal of Vision*, *9*, 1–16. doi:10.1167/9.2.7

Bodamer, J. (1947). Die Prosop-Agnosie (Die Agnosie des Physiognomieerkennens). *Archiv fur Psychiatrie und Nervenkrankheiten*, *179*, 6–54. doi:10.1007/BF00352849

Boutsen, L., Humphreys, G. W., Praamstra, P., & Warbrick, T. (2006). Comparing neural correlates of configural processing in faces and objects: An ERP study of the Thatcher illusion. *Neuroimage*, *32*, 352–367. doi:10.1016/j.neuroimage.2006.03.023

Bowles, D. C., McKone, E., Dawel, A., Duchaine, B., Palermo, R., Schmalzl, L., . . . Yovel, G. (2009). Diagnosing prosopagnosia: Effects of ageing, sex, and participant-stimulus ethnic match on the Cambridge Face Memory Test and Cambridge Face Perception Test. *Cognitive Neuropsychology*, *26*, 423–455. doi:10.1080/02643290903343149

Bruce, V., & Young, A. W. (2012). *Face perception*. London: Psychology Press.

Bruyer, R. (2011). Configural face processing: A meta-analytic survey. *Perception*, *40*, 1478–1490. doi:10.1068/p6928

Bukach, C. M., Bub, D. M., Gauthier, I., & Tarr, M. J. (2006). Perceptual expertise effects are not all or none: Local perceptual expertise for faces in a case of prosopagnosia. *Journal of Cognitive Neuroscience*, *18*, 48–63. doi:10.1162/089892906775250094

Bukach, C. M., Gauthier, I., Tarr, M. J., Kadlec, H., Barth, S., Ryan, E., . . . Bub, D.N. (2012). Does acquisition of Greeble expertise in prosopagnosia rule out a domain–general deficit? *Neuropsychologia*, *50*, 289–304. doi:10.1016/j.neuropsychologia.2011.11.023

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

Busigny, T., Joubert, S., Felician, O., Ceccaldi, M., & Rossion, B. (2010). Holistic perception of the individual face is specific and necessary: Evidence from an extensive case study of acquired prosopagnosia. *Neuropsychologia*, *48*, 4057–4092. doi:10.1016/j.neuropsychologia.2010.09.017

Busigny, T., Van Belle, G., Jemel, B., Hosein, A., Joubert, S., Bub, D., Tanaka, J.W., & Rossion, B. (submitted). Face-specific impairment in holistic perception following focal lesion to the right anterior temporal lobe.

Busigny, T., Van Belle, G., Jemel, B., Hosein, A., & Rossion, B. (2012). The neuropsychology of the right anterior temporal lobe: A perceptual involvement in face recognition. *Manuscript in preparation*.

Cabeza, R., & Kato, T. (2000). Features are also important: Contributions of featural and configural processing to face recognition. *Psychological Science*, *11*, 419–433. doi:10.1111/1467-9280.00283

Caldara, R., Schyns, P., Mayer, E., Smith, M., Gosselin, F., & Rossion, B. (2005). Does prosopagnosia take the eyes out from faces? Evidence for a defect in the use of diagnostic facial information in a brain-damaged patient. *Journal of Cognitive Neuroscience*, *17*, 1652–1666. doi:10.1162/089892905774597254

Calder, A. J., Young, A. W., Keane, J., & Dean, M. J. (2000). Configural information in facial expression perception. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 527–551. doi:10.1037/0096-1523.26.2.527

Carbon, C. C., Schweinberger, S. R., Kaufmann, J. M., & Leder, H. (2005). The Thatcher illusion seen by the brain: An event-related brain potentials study. *Brain Research: Cognitive Brain Research*, *24*, 544–555. doi:10.1016/j.cogbrainres.2005.03.008

Carbon, C. C., Grüter, T., Weber, J. E., & Lueschow, A. (2007). Faces as objects of non-expertise: processing of thatcherised faces in congenital prosopagnosia. *Perception*, *36*, 1635–1645. doi:10.1068/p5467

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

Carey, S., & Diamond, R. (1977). From piecemeal to configurational representation of faces. *Science*, *195*, 312–314. doi:10.1126/science.831281

Carey, S., & Diamond, R. (1994). Are faces perceived as configurations more by adults than by children? *Visual Cognition*, *1*, 253–274. doi:10.1080/13506289408402302

Chaudhuri, A., & Albright, T. D. (1997). Neuronal responses to edges defined by luminance vs. temporal texture in macaque area V1. *Visual Neuroscience*, *14*, 949–962. doi:10.1017/S0952523800011664

Chen, Z. (2012). Object-based attention: A tutorial review. *Attention, Perception, and Psychophysics*, *74*, 784–802. doi:10.3758/s13414-012-0322-z

Cheung, O. S., Richler, J. J., Palmeri, T. J., & Gauthier, I. (2008). Revisiting the role of spatial frequencies in the holistic processing of faces. *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 1327–1336. doi:10.1037/a0011752

Cornes, K., Donnelly, N., Godwin, H., & Wenger, M.J. (2011). Perceptual and decisional factors influencing the discrimination of inversion in the Thatcher illusion. *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 645–668. doi:10.1037/a0020985

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

Curby, K. M., Goldstein, R. R., & Blacker, K. (2013). Disrupting perceptual grouping of face parts impairs holistic face processing. *Attention, Perception, and Psychophysics*, *75*(1), 83–91. doi:10.3758/s13414-012-0386-9

Dahl, C. D., Logothetis, N. K., & Hoffman, K. L. (2007). Individuation and holistic processing of faces in rhesus monkeys. *Proceedings of the Royal Society: Biological Sciences*, *274B*, 2069–2076. doi:10.1098/rspb.2007.0477

Dalrymple-Alford, E. C., & Budayr, B. (1966). Examination of some aspects of the Stroop Color-Word Test. *Perceptual and Motor Skills*, *23*, 1211–1214. doi:10.2466/pms.1966.23.3f.1211

Davidoff, J., & Donnelly, N. (1990). Object superiority effects: Complete versus part probes. *Acta Psychologica*, *73*, 225–243. doi:10.1016/0001-6918(90)90024-A

Davies, G., Ellis, H., & Shepherd, J. (1977). Cue saliency in faces as assessed by the "photofit" technique. *Perception*, *6*, 263–269. doi:10.1068/p060263

DeGutis, J., Wilmer, J., Mercado, R. J., & Cohan, S. (2013). Using regression to measure holistic face processing reveals a strong link with face recognition ability. *Cognition*, *126*, 87–100. doi:10.1016/j.cognition.2012.09.004

de Haan, E. H. F., Young, A. W., & Newcombe, F. (1987). Face recognition without awareness. *Cognitive Neuropsychology*, *4*, 385–415. doi:10.1080/02643298708252045

de Heering, A., Aljuhanay, A., Rossion, B., & Pascalis, O. (2012). Early deafness increases the face inversion effect but does not modulate the composite face effect. *Frontiers in Psychology*, *3*, 124. doi:10.3389/fpsyg.2012.00124

de Heering, A., Houthuys, S., & Rossion, B. (2007). Holistic face processing is mature at 4 years of age: Evidence from the composite face effect. *Journal of Experimental Child Psychology*, *96*, 57–70. doi:10.1016/j.jecp.2006.07.001

de Heering, A., & Maurer, D. (2012). Face memory deficits in patients deprived of early visual input by bilateral congenital cataracts. *Developmental Psychobiology*. unpublished

de Heering, A., & Rossion, B. (2008). Prolonged visual experience in adulthood modulates holistic face perception. *PLOS One*, *3*, e2317. doi:10.1371/journal.pone.0002317

de Heering, A., Rossion, B., Turati, C., & Simion, F. (2008). Holistic face processing can be independent of gaze behaviour: Evidence from the face composite effect. *Journal of Neuropsychology*, *2*, 183–195. doi:10.1348/174866407X251694

de Heering, A., Wallis, J., & Maurer, D. (2012). The composite face effect survives asymmetric face distortions. *Perception*, *41*, 707–716. doi:10.1068/p7212

Dennett, H. W., McKone, E., Tavashmi, R., Hall, A., Pidcock, M., Edwards, M., & Duchaine, B. (2011). The Cambridge Car Memory Test: A task matched in format to the Cambridge Face Memory Test, with norms, reliability, sex differences, dissociations from face memory, and expertise effects. *Behavior Research Methods*, *44*, 587–605. doi:10.3758/s13428-011-0160-?2

Donnelly, N., & Davidoff, J. (1999). The mental representations of faces and houses: Issues concerning parts and wholes. *Visual Cognition*, *6*, 319–343. doi:10.1080/135062899395000

Duchaine, B. C., & Nakayama, K. (2006a). Developmental prosopagnosia: A window to content-specific face processing. *Current Opinions in Neurobiology*, *16*, 166–173. doi:10.1016/j.conb.2006.03.003

Duchaine, B. C., & Nakayama, K. (2006b). The Cambridge Face Memory Test: Results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. *Neuropsychologia*, *44*, 576–585. doi:10.1016/j.neuropsychologia.2005.07.001

Edmonds, A. J., & Lewis, M. B. (2007). The effect of rotation on configural encoding in a face-matching task. *Perception*, *36*, 446–460. doi:10.1068/p5530

Ellis, H. D., & Florence, M. (1990). Bodamer's (1947) paper on prosopagnosia. *Cognitive Neuropsychology*, *7*, 81–105. doi:10.1080/02643299008253437

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception and Psychophysics*, *16*, 143–149. doi:10.3758/BF03203267

Eriksen, C. W. (1995). The Flankers Task and response competition: A useful tool for investigating a variety of cognitive problems. *Visual Cognition*, *2*, 101–118. doi:10.1080/13506289508401726

Farah, M. J., Klein, K. L., & Levinson, K. L. (1995). Face perception and within-category discrimination in prosopagnosia. *Neuropsychologia*, *33*, 661–674. doi:10.1016/0028-3932(95)00002-K

Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. W. (1998). What is "special" about face perception? *Psychological Review*, *105*, 482–498. doi:10.1037/0033-295X.105.3.482

Freire, A., Lee, K., & Symons, L. A. (2000). The face-inversion effect as a deficit in the encoding of configural information: Direct evidence. *Perception*, *29*, 159–170. doi:10.1068/p3012

Freiwald, W. A., Tsao, D. Y., & Livingstone, M. S. (2009). A face feature space in the macaque temporal lobe. *Nature Neuroscience*, *12*, 1187–1196. doi:10.1038/nn.2363

Galton, F. (1883). *Inquiries into human faculty and its development*. London: Macmillan.

Gao, Z., Flevaris, A. V., Robertson, L. C., & Bentin, S. (2011). Priming global and local processing of composite faces: Revisiting the processing-bias effect on face perception. *Attention, Perception, and Psychophysics*, *73*, 1477–1486. doi:10.3758/s13414-011-0109-7

Garner, W. R. (1974). *The processing of information and structure*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Garner, W. R. (1988). Facilitation and interference with a separable redundant dimension in stimulus comparison. *Perception & Psychophysics*, *44*, 321–330. doi:10.3758/BF03210413

Gauthier, I., & Bukach, C. M. (2007). Should we reject the expertise hypothesis? *Cognition*, *103*, 322–330. doi:10.1016/j.cognition.2006.05.003

Gauthier, I., Curran, T., Curby, K. M., & Collins, D. (2003). Perceptual interference supports a non-modular account of face processing. *Nature Neuroscience*, *6*, 428–432. doi:10.1038/nn1029

Gauthier, I., Klaiman, C., & Schultz, R. T. (2009). Face composite effects reveal abnormal face processing in Autism Spectrum Disorders. *Vision Research*, *49*, 470–478. doi:10.1016/j.visres.2008.12.007

Gauthier, I., & Tarr, M. J. (2002). Unraveling mechanisms for expert object recognition: Bridging brain activity and behaviour. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 431–466. doi:10.1037/0096-1523.28.2.431

Gauthier, I., Williams, P., Tarr, M. J., & Tanaka, J. W. (1998). Training "greeble" experts: A framework for studying expert object recognition processes. *Vision Research*, *38*, 2401–2428. doi:10.1016/S0042-6989(97)00442-2

Germine, L. T., Duchaine, B., & Nakayama, K. (2011). Where cognitive development and aging meet: Face learning ability peaks after age 30. *Cognition*, *118*, 201–210. doi:10.1016/j.cognition.2010.11.002

Ginsburg, A. P. (1978). *Visual information processing based on spatial filters constrained by biological data* (Unpublished doctoral dissertation). Cambridge University, Cambridge, UK.

Goffaux, V. (2009). Spatial interactions in upright and inverted faces: Re-exploration of spatial scale influence. *Vision Research*, *49*, 774–781. doi:10.1016/j.visres.2009.02.009

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

Goffaux, V., & Rossion, B. (2006). Faces are "spatial"—Holistic face perception is supported by low spatial frequencies. *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 1023–1039. doi:10.1037/0096-1523.32.4.1023

Goffaux, V., & Rossion, B. (2007). Face inversion disproportionately impairs the perception of vertical but not horizontal relations between features. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 995–1002. doi:10.1037/0096-1523.33.4.995

Gold, J., Mundy, P., & Tjan, B. (2012). The perception of a face is no more than the sum of its parts. *Psychological Science*, *23*, 427–434. doi:10.1177/0956797611427407

Goldfarb, L., & Henik, A. (2006). New data analysis of the Stroop matching task calls for a reevaluation of theory. *Psychological Science*, *17*, 96–100. doi:10.1111/j.1467-9280.2006.01670.x

Goldstein, A. G., & Chance, J. E. (1980). Memory for faces and schema theory. *Journal of Psychology*, *105*, 47–59. doi:10.1080/00223980.1980.9915131

Gorea, A., & Julesz, B. (1990). Context superiority in a detection task with line-element stimuli: A low-level effect. *Perception*, *19*, 5–16. doi:10.1068/p190005

Gosselin, F., & Schyns, P. G. (2001). Bubbles: A technique to reveal the use of information in recognition tasks. *Vision Research*, *41*, 2261–2271. doi:10.1016/S0042-6989(01)00097-9

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York, NY: Wiley.

Gregory, R. L. (1997). Knowledge in perception and illusion. *Philosophical Transactions of the Royal Society of London*, *352B*, 1121–1128. doi:10.1098/rstb.1997.0095

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: Neural models of stimulus-specific effects. *Trends in Cognitive Science*, *10*, 14–23. doi:10.1016/j.tics.2005.11.006

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

Haig, N. D. (1984). The effect of feature displacement on face recognition. *Perception*, *13*, 502–512. doi:10.1068/p130505

Haig, N. D. (1985). How faces differ: A new comparative technique. *Perception*, *14*, 601–615. doi:10.1068/p140601

Haig, N. D. (1986). Investigating face recognition with an image processing computer. In H. D. Ellis, M. A. Jeeves, F. Newcombe, & A. W. Young (Eds.), *Aspects of face processing* (pp. 410–425). Dordrecht: Martinus Nijhof.

Harris, A., & Aguirre, G. K. (2008). The representation of parts and wholes in face-selective cortex. *Journal of Cognitive Neuroscience*, *20*, 863–878. doi:10.1162/jocn.2008.20509

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

Hochberg, J., & Galper, R. E. (1967). Recognition of faces: I. An exploratory study. *Psychonomic Science*, *9*, 619–620.

Hole, G., & George, P. (2011). Evidence for holistic processing of facial age. *Visual Cognition*, *19*, 585–615. doi:10.1080/13506285.2011.562076

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

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

Homa, D., Haver, B., & Schwartz, T. (1976). Perceptibility of schematic face stimuli: Evidence for a perceptual Gestalt. *Memory and Cognition*, *4*, 176–185. doi:10.3758/BF03213162

Hommel, B. (1997). Interactions between stimulus-stimulus congruence and stimulus-response compatibility. *Psychological Research*, *59*, 248–260. doi:10.1007/BF00439302

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

Hsiao, J. H., & Cottrell, G. W. (2009). Not all visual expertise is holistic, but it may be leftist: The case of Chinese character recognition. *Psychological Science*, *20*, 455–463. doi:10.1111/j.1467-9280.2009.02315.x

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

Hugenberg, K., & Corneille, O. (2009). Holistic processing is tuned for in-group faces. *Cognitive Science*, *33*, 1173–1181. doi:10.1111/j.1551-6709.2009.01048.x

Hulten, P. (1987). *The Arcimboldo effect*. London: Thames & Hudson.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, *40*, 1489–1506. doi:10.1016/S0042-6989(99)00163-7

Jacques, C., d'Arripe, O., & Rossion, B. (2007). The time course of the face inversion effect during individual face discrimination. *Journal of Vision*, *7*, 1–9. doi:10.1167/7.8.3

Jacques, C., & Rossion, B. (2009). The initial representation of individual faces in the right occipito-temporal cortex is holistic: Electrophysiological evidence from the composite face illusion. *Journal of Vision*, *11*(9), 1–16.

Jacques, C., & Rossion, B. (2010). Misaligning face halves increases and delays the N170 specifically for upright faces: Implications for the nature of early face representations. *Brain Research*, *1318*, 96–109. doi:10.1016/j.brainres.2009.12.070

Jacobs, A. M. (1986). Eye-movement control in visual search: How direct is visual span control? *Perception and Psychophysics*, *39*, 47–58. doi:10.3758/BF03207583

Jiang, F., Blanz, V., & Rossion, B. (2011). Holistic processing of shape cues in face identification: Evidence from face inversion, composite faces and acquired prosopagnosia. *Visual Cognition*, *19*, 1003–1034. doi:10.1080/13506285.2011.604360

Jung, R., & Spillmann, L. (1970). Receptive-field estimation and perceptual integration in human vision. In F. A. Young & D. B. Lindsley (Eds.), *Early experience and visual information processing in perceptual and reading disorders* (pp. 181–197). Washington, DC: National Academic Press.

Jung, R., von Baumgarten, R., & Baumgartner, G. (1952). Mikroableitungen von einzelneç n Nervenzellen im optischen Cortex der Katze: Die lichtaktivierten B-Neurone. *Archiv fur Psychiatrie und Nervenkrankheiten*, *189*, 521–539. doi:10.1007/BF00394865

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.

Kastner, S., De Weerd, P., Desimone, R., & Ungerleider, L. G. (2000). Texture segregation in human extrastriate cortex: A fMRI study. *Journal of Neurophysiology*, *83*, 2453–2457.

Kelly, K. R., Gallie, B. L., & Steeves, J. K. (2012). Impaired face processing in early monocular deprivation from enucleation. *Optometry and Vision Science*, *89*, 137–147. doi:10.1097/OPX.0b013e318240488e

Khurana, B., Carter, R. M., Watanabe, K., & Nijhawan, R. (2006). Flash-lag chimeras: The role of perceived alignment in the composite face effect. *Vision Research*, *46*, 2757–2772. doi:10.1016/j.visres.2006.02.001

Kimchi, R. (1992). Primacy of wholistic processing and global/local paradigm: A critical review. *Psychological Bulletin*, *112*, 24–38. doi:10.1037/0033-2909.112.1.24

Kimchi, R. (2009). Perceptual organization and visual attention. *Progress in Brain Research*, *176*, 15–33.

Kimchi, R., Behrmann, M., & Olson, C. R. (Eds.). (2003). *Perceptual organization in vision: Behavioral and neural perspectives.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Archives of General Psychiatry*, *59*, 809–816. doi:10.1001/archpsyc.59.9.809

Kobatake, E., & Tanaka, K. (1994). Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *Journal of Neurophysiology*, *71*, 856–867.

Koffka, K. (1963). *Principles of Gestalt psychology.* New York, NY: Harcourt Brace Jovanovich. (Original work published 1935)

Köhler, W. (1929). *Gestalt psychology.* New York, NY: Liveright.

Konar, Y., Bennett, P. J., & Sekuler, A. B. (2010). Holistic processing is not correlated with face-identification accuracy. *Psychological Science*, *21*, 38–43. doi:10.1177/0956797609356508

Kramer, A. F., & Jacobson, A. (1991). Perceptual organization and focused attention: The role of objects and proximity in visual processing. *Perception and Psychophysics*, *50*, 267–284. doi:10.3758/BF03206750

Kuefner, D., Jacques, C., Prieto, E. A., & Rossion, B. (2010). Electrophysiological correlates of the composite face illusion: Disentangling perceptual and decisional components of holistic face processing in the human brain. *Brain and Cognition*, *74*, 225–238. doi:10.1016/j.bandc.2010.08.001

Kuefner, D., Macchi Cassia, V., Vescovo, E., & Picozzi, M. (2010). Natural experience acquired in adulthood enhances holistic processing of other-age faces. *Visual Cognition*, *18*, 11–25. doi:10.1080/13506280802396507

Laguesse, R., & Rossion, B. (2011). The composite face illusion and its disappearance with misaligned faces: An effect of metric distance or part separation? [Abstract]. *Journal of Vision*, *11*(11), 617. doi:10.1167/11.11.617

Laurence, S., & Hole, G. (2012). Identity specific adaptation with composite faces. *Visual Cognition*, *20*, 109–120. doi:10.1080/13506285.2012.655805

Leder, H., & Bruce, V. (2000). When inverted faces are recognized: The role of configural information in face recognition. *Quarterly Journal of Experimental Psychology*, *53A*, 513–536.

Lee, K., & Freire, A. (1999). Effects of face configuration change on shape perception: A new illusion. *Perception*, *28*, 1217–1226. doi:10.1068/p2972

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

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

Letourneau, S. M., & Mitchell, T. V. (2008). Behavioural and ERP measures of holistic face processing in a composite task. *Brain and Cognition*, *67*, 234–245. doi:10.1016/j.bandc.2008.01.007

Levine, D. N., & Calvanio, R. (1989). Prosopagnosia: A defect in visual configural processing. *Brain and Cognition*, *10*, 149–147. doi:10.1016/0278-2626(89)90051-1

Lewis, M. B. (2001). The Lady's not for turning: Rotation of the Thatcher illusion. *Perception*, *30*, 769–774. doi:10.1068/p3174

Lewis, M. B., & Johnston, R. A. (1997). The Thatcher illusion as a test of configural disruption. *Perception*, *26*, 225–227. doi:10.1068/p260225

Lowe, D. G. (1985). *Perceptual organization and visual recognition*. Boston, MA: Kluwer Academic Publishers.

Luck, S. J. (2005). *An Introduction to the Event-Related Potential Technique*. Cambridge, MA: MIT Press.

Luo, C. R. (1999). Semantic competition as the basis of Stroop interference: Evidence from color–word matching tasks. *Psychological Science*, *10*, 35–40. doi:10.1111/1467-9280.00103

Macchi Cassia, V., Picozzi, M., Kuefner, D., Bricolo, E., & Turati, C. (2009). Holistic processing for faces and cars in preschool-aged children and adults: Evidence from the composite effect. *Developmental Science*, *12*, 236–248. doi:10.1111/j.1467-7687.2008.00765.x

Mach, E. (1865). Über die Wirkung der räummlichen Verthelung des Lich- treizes auf die Netzhaut. *Akad Wiss Lit Abh Math-Natwiss KL (Mainz)*, *52*, 303–322.

Mack, M. L., Richler, J. J., Gauthier, I., & Palmeri, T. J. (2011). Indecision on decisional separability. *Psychonomic Bulletin and Review*, *18*, 1–9. doi:10.3758/s13423-010-0017-1

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, *109*, 163–203. doi:10.1037/0033-2909.109.2.163

Malpass, R. S., & Kravitz, J. (1969). Recognition for faces of own and other race. *Journal of Personality and Social Psychology*, *13*, 330–334. doi:10.1037/h0028434

Martini, P., McKone, E., & Nakayama, K. (2001). Orientation tuning of human face processing estimated by contrast matching in transparency displays. *Journal of Vision*, *1*, 335.

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*, *6*, 255–260. doi:10.1016/S1364-6613(02)01903-4

McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception and Psychophysics*, *17*, 578–586. doi:10.3758/BF03203972

McKone, E. (2004). Isolating the special component of face recognition: Peripheral identification and a Mooney face. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 181–197. doi:10.1037/0278-7393.30.1.181

McKone, E. (2008). Configural processing and face viewpoint. *Experimental Psychology: Human Perception and Performance*, *34*, 310–327. doi:10.1037/0096-1523.34.2.310

McKone, E. (2009). Holistic processing for faces operates over a wide range of sizes but is strongest at identification rather than conversational distances. *Vision Research*, *49*, 268–283. doi:10.1016/j.visres.2008.10.020

McKone, E., Martini, P., & Nakayama, K. (2001). Categorical perception of face identity in noise isolates configural processing. *Journal of experimental psychology: Human perception and performance*, *27*, 573–599.

McKone, E., Martini, P., & Nakayama, K. (2003). Isolating holistic processing in faces (and perhaps objects). In M. A. Peterson & G. Rhodes (Eds.), *Perception of faces, objects, and scenes: Analytic and holistic processes* (92–119). New York: Oxford University Press.

McKone, E., & Robbins, R. (2007). The evidence rejects the expertise hypothesis: Reply to Gauthier & Bukach. *Cognition*, *103*, 331–336. doi:10.1016/j.cognition.2006.05.014

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

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 and Review*, *16*, 778–797. doi:10.3758/PBR.16.5.778

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

Meinhardt-Injac, B., Persike, M., & Meinhardt, G. (2010). The time course of face matching by internal and external features: Effects of context and inversion. *Vision Research*, *50*, 1598–1611. doi:10.1016/j.visres.2010.05.018

Meinhardt-Injac, B., Persike, M., & Meinhardt, G. (2011). The context effect in face matching: Effects of feedback. *Vision Research*, *51*, 2121–2131. doi:10.1016/j.visres.2011.08.004

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analysis review. *Psychology, Public Policy, and Law*, *7*, 3–35. doi:10.1037/1076-8971.7.1.3

Mermelstein, R., Banks, W., & Prinzmetal, W. (1979). Figural goodness effects in perception and memory. *Perception and Psychophysics*, *26*, 472–480. doi:10.3758/BF03204287

Michel, C., Caldara, R., & Rossion, B. (2006). Same-race faces are perceived more holistically than other-race faces. *Visual Cognition*, *14*, 55–73. doi:10.1080/13506280500158761

Michel, C., Corneille, O., & Rossion, B. (2007). Race categorization modulates holistic encoding. *Cognitive Science*, *31*, 911–924. doi:10.1080/03640210701530805

Michel, C., Corneille, O., & Rossion, B. (2010). Holistic face encoding is modulated by perceived face race: Evidence from perceptual adaptation. *Visual Cognition*, *18*, 434–455. doi:10.1080/13506280902819697

Michel, C., Rossion, B., Han, J., Chung, C. H., & Caldara, R. (2006). Holistic processing is finely tuned for faces of one's own race. *Psychological Science*, *17*, 608–615. doi:10.1111/j.1467-9280.2006.01752.x

Miller, J. (1991). The flanker compatibility effect as a function of visual angle, attentional focus, visual transients, and perceptual load: A search for boundary conditions. *Perception and Psychophysics*, *49*, 270–288. doi:10.3758/BF03214311

Mondloch, C. J., Elms, N., Maurer, D., Rhodes, G., Hayward, W. G., Tanaka, J. W., & Zhou, G. (2010). Processes underlying the cross-race effect: An investigation of holistic, featural, and relational processing of own-race versus other-race faces. *Perception*, *39*, 1065–1085. doi:10.1068/p6608

Mondloch, C. J., & Maurer, D. (2008). The effect of face orientation on holistic processing. *Perception*, *37*, 1175–1186. doi:10.1068/p6048

Mondloch, C. J., Pathman, T., Maurer, D., Le Grand, R., & de Schonen, S. (2007). The composite face effect in six-year-old children: Evidence for adult-like holistic processing. *Visual Cognition*, *15*, 564–577. doi:10.1080/13506280600859383

Mooney, C. M. (1957). Age in the development of closure ability in children. *Canadian Journal of Psychology*, *11*, 219–226. doi:10.1037/h0083717

Moore, C., & Cavanagh, P. (1998). Recovery of 3D volume from 2-tone images of novel objects. *Cognition*, *67*, 45–71. doi:10.1016/S0010-0277(98)00014-6

Morikawa, K., Okumura, K., & Matsushita, S. (2012). Head size illusion: Head outlines are processed holistically too. *Perception*, *41*, 5. (ECVP Abstract Suppl.).

Morrison, D. J., & Schyns, P. G. (2001). Usage of spatial scales for the categorization of faces, objects, and scenes. *Psychonomic Bulletin and Review*, *8*, 454–469. doi:10.3758/BF03196180

Müller-Lyer, F. C. (1889). Optische Urteilstäuschungen. *Archiv für Physiologie*, (Suppl.), 263–270.

Murray, J., Yong, E., & Rhodes, G. (2000). Revisiting the perception of upside-down faces. *Psychological Science*, *11*, 492–496. doi:10.1111/1467-9280.00294

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, *9*, 353–383. doi:10.1016/0010-0285(77)90012-3

Navon, D. (2003). What does a compound letter tell the psychologist's mind? *Acta Psychologica*, *114*, 273–309. doi:10.1016/j.actpsy.2003.06.002

Neri, P., & Levi, D. M. (2006). Receptive versus perceptive fields from the reverse-correlation viewpoint. *Vision Research*, *46*, 2465–2474. doi:10.1016/j.visres.2006.02.002

Nishida, S., & Johnston, A. (1999). Influence of motion signals on the perceived position of spatial pattern. *Nature*, *397*, 610–612. doi:10.1038/17600

Nishimura, M., Rutherford, M. D., & Maurer, D. (2008). Converging evidence of configural processing of faces in high-functioning adults with autism spectrum disorders. *Visual Cognition*, *7*, 859–891. doi:10.1080/13506280701538514

Nothdurft, H. C., Gallant, J. L., & Van Essen, D. C. (2000). Response profiles to texture border patterns in area V1. *Visual Neuroscience*, *17*, 421–436. doi:10.1017/S0952523800173092

Orban de Xivry, J.-J., Ramon, M., Lefèvre, P., & Rossion, B. (2008). Reduced fixation on the upper area of personally familiar faces following acquired prosopagnosia. *Journal of Neuropsychology*, *2*, 245–268. doi:10.1348/174866407X260199

O'Toole, A. J. (2011). Cognitive and computational approaches to face perception. In A. Calder, G. Rhodes, M. Johnson, & J. V. Haxby (Eds.), *Oxford handbook of face perception* (pp. 15–30). Oxford: Oxford University Press.

Palermo, R., Willis, M. L., Rivolta, D., McKone, E., Wilson, C. E., & Calder, A. (2011). Impaired holistic coding of facial expression and facial identity in congenital prosopagnosia. *Neuropsychologia*, *49*, 1226–1235. doi:10.1016/j.neuropsychologia.2011.02.021

Pellicano, E., & Rhodes, G. (2003). Holistic processing of faces in preschool children and adults. *Psychological Science*, *14*, 618–622. doi:10.1046/j.0956-7976.2003.psci_1474.x

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

Pomerantz, J. R., Agrawal, A., Jewell, S. W., Jeong, M., Khan, H., & Lozano, S. C. (2003). Contour grouping inside and outside of facial contexts. *Acta Psychologica*, *114*, 245–271. doi:10.1016/j.actpsy.2003.08.004

Pomerantz, J. R., Carson, C. E., & Feldman, E. M. (1994). Interference effects in perceptual organization. In S. Ballesteros (Ed.), *Cognitive approaches to human perception* (pp. 123–152). Hillsdale, NJ: Lawrence Erlbaum Associates.

Pomerantz, J. R., & Kubovy, M. (1986). Theoretical approaches to perceptual organization. In K. R. Boff, L. Kaufman, & J. Thomas (Eds.), *Handbook of perception and human performance* (Vol. 3, pp. 3601–3646). New York, NY: Wiley.

Pomerantz, J. R., & Portillo, M. C. (2011). Grouping and emergent features in vision: Toward a theory of basic Gestalts. *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 1331–1349.

Pomerantz, J. R., & Pristach, E. A. (1989). Emergent features, attention, and perceptual glue in visual form perception. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 635–649. doi:10.1037/0096-1523.15.4.635

Pomerantz, J. R., Pristach, E. A., & Carson, C. E. (1989). Attention and object perception. In B. Shepp & S. Ballesteros (Eds.), *Object perception: Structure and process* (pp. 53–89). Hillsdale, NJ: Lawrence Erlbaum Associates.

Pomerantz, J. R., Sager, L. C., & Stoever, R. J. (1977). Perception of wholes and of their component parts: Some configural superiority effects. *Journal of Experimental Psychology: Human Perception and Performance*, *3*, 422–435. doi:10.1037/0096-1523.3.3.422

Rakover, S. S. (2012). A feature-inversion effect: Can an isolated feature show behaviour like the face-inversion effect? *Psychonomic Bulletin and Review*, *19*(4), 617–624. doi:10.3758/s13423-012-0264-4

Rakover, S. S., & Teucher, B. (1997). Facial inversion effects: Parts and whole relationship. *Perception and Psychophysics*, *59*, 752–761. doi:10.3758/BF03206021

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

Ramon, M., & Rossion, B. (2012). Hemisphere-dependent holistic processing of familiar faces. *Brain and Cognition*, *78*, 7–13.

Ratliff, F. (1965). *Mach bands: Quantitative studies on neural networks in the retina*. San Francisco, CA: Holden-Day.

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology*, *7*, 65–81. doi:10.1016/0010-0285(75)90005-5

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*, 372–422. doi:10.1037/0033-2909.124.3.372

Regan, D. (2000). *Human perception of objects: Early visual processing of spatial form defined by luminance, color, texture, motion, and binocular disparity*. Sunderland, MA: Sinauer Associates.

Reingold, E. M., Charness, N., Pomplun, M., & Stampe, D. M. (2001). Visual span in expert chess players: Evidence from eye movements. *Psychological Science*, *12*, 49–56. doi:10.1111/1467-9280.00309

Rezlescu, C., Pitcher, D., & Duchaine, B. (2012). Acquired prosopagnosia with spared within-class object recognition but impaired recognition of degraded basic-level objects. *Cognitive Neuropsychology*, *29*, 325–347. doi:10.1080/02643294.2012.749223

Rhodes, G. (1988). Looking at faces: First-order and second-order features as determinates of facial appearance. *Perception*, *17*, 43–63. doi:10.1068/p170043

Rhodes, G., Brake, S., & Atkinson, A. P. (1993). What's lost in inverted faces? *Cognition*, *47*, 25–57. doi:10.1016/0010-0277(93)90061-Y

Richler, J. J., Cheung, O. S., & Gauthier, I. (2011a). Beliefs alter holistic face processing . . . if response bias is not taken into account. *Journal of Vision*, *11*, 13. doi:10.1167/11.13.17

Richler, J. J., Cheung, O. S., & Gauthier, I. (2011b). Holistic processing predicts face recognition. *Psychological Science*, *20*, 464–471.

Richler, J. J., Cheung, O. S., Wong, A. C.-N., & Gauthier, I. (2009). Does response interference contribute to face composite effects? *Psychonomic Bulletin and Review*, *16*, 258–263. doi:10.3758/PBR.16.2.258

Richler, J. J., Gauthier, I., Wenger, M. J., & Palmeri, T. J. (2008). Holistic processing of faces: Perceptual and decisional components. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 328–342. doi:10.1037/0278-7393.34.2.328

Richler, J. J., Mack, M. L., Gauthier, I., & Palmeri, T. J. (2009). Holistic processing happens at a glance. *Vision Research*, *49*, 2856–2861. doi:10.1016/j.visres.2009.08.025

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

Richler, J. J., Tanaka, J. W., Brown, D. D., & Gauthier, I. (2008). Why does selective attention to parts fail in face processing? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 1356–1368. doi:10.1037/a0013080

Richler, J. J., Wong, Y. K., & Gauthier, I. (2011). Perceptual expertise as a shift from strategic interference to automatic holistic processing. *Current Directions in Psychological Science*, *20*, 129–134. doi:10.1177/0963721411402472

Robbins, R., & McKone, E. (2003). Can holistic processing be learned for inverted faces? *Cognition*, *88*, 79–107. doi:10.1016/S0010-0277(03)00020-9

Robbins, R., & McKone, E. (2007). No face-like processing for objects-of-expertise in three behavioural tasks. *Cognition*, *103*, 34–79. doi:10.1016/j.cognition.2006.02.008

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

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. (2009). Distinguishing the cause and consequence of face inversion: The perceptual field hypothesis. *Acta Psychologica*, *132*, 300–312. doi:10.1016/j.actpsy.2009.08.002

Rossion, B., & Boremanse, A. (2008). Nonlinear relationship between holistic processing of individual faces and picture-plane rotation: Evidence from the face composite illusion. *Journal of Vision*, *8*(4), 1–13. doi:10.1167/8.4.3

Rossion, B., Caldara, R., Seghier, M., Schuller, A.-M., Lazeyras, F., & Mayer, E. (2003). A network of occipito-temporal face-sensitive areas besides the right middle fusiform gyrus is necessary for normal face processing. *Brain*, *126*, 2381–2395. doi:10.1093/brain/awg241

Rossion, B., Dricot, L., Goebel, R., & Busigny, T. (2011). Holistic face categorization in higher-level cortical visual areas of the normal and prosopagnosic brain: Towards a non-hierarchical view of face perception. *Frontiers in Human Neuroscience*, *4*, 225. doi:10.3389/fnhum.2010.00225

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. (2008). Does physical interstimulus variance account for early electrophysiological face sensitive responses in the human brain? Ten lessons on the N170. *NeuroImage*, *39*, 1959–1979. doi:10.1016/j.neuroimage.2007.10.011

Rossion, B., & Jacques, C. (2011). The N170: Understanding the time-course of face perception in the human brain. In S. Luck & E. Kappenman (Eds.), *The Oxford handbook of ERP components* (pp. 115–142). Oxford: Oxford University Press.

Rossion, B., & Michel, C. (2011). An experienced-based holistic account of the other-race face effect. In A. Calder, G. Rhodes, J. V. Haxby, & M. Johnson (Eds.), *The Oxford handbook of face perception* (pp. 215–244). Oxford: Oxford University Press.

Rossion, B., Kaiser, M. D., Bub, D., & Tanaka, J. W. (2009). Is the loss of diagnosticity of the eye region a common feature of acquired prosopagnosia? *Journal of Neuropsychology*, *3*, 69–78. doi:10.1348/174866408X289944

Russell, R., Duchaine, B., & Nakayama, K. (2009). Super-recognizers: People with extraordinary face recognition ability. *Psychonomic Bulletin and Review*, *16*, 252–257. doi:10.3758/PBR.16.2.252

Sadr, J., Jarudi, I., & Sinha, P. (2003). The role of eyebrows in face recognition. *Perception*, *32*, 285–293. doi:10.1068/p5027

Schiltz, C., Dricot, L., Goebel, R., & Rossion, B. (2010). Holistic perception of individual faces in the right middle fusiform gyrus as evidenced by the composite face illusion. *Journal of Vision*, *10*(2), 1–16. doi:10.1167/10.2.25

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

Schmalzl, L., Palermo, R., & Coltheart, M. (2008). Cognitive heterogeneity in genetically based prosopagnosia: A family study. *Journal of Neuropsychology*, *2*, 99–117. doi:10.1348/174866407X256554

Schyns, P. G., Jentzsch, I., Johnson, M., Schweinberger, S. R., & Gosselin, F. (2003). A principled method for determining the functionality of brain responses. *Neuroreport*, *14*, 1665–1669. doi:10.1097/00001756-200309150-00002

Seitz, K. (2002). Parts and wholes in person recognition: Developmental trends. *Journal of Experimental Child Psychology*, *82*, 367–381. doi:10.1016/S0022-0965(02)00106-6

Sekuler, A. B., Gaspar, C. M., Gold, J. M., & Bennett, P. J. (2004). Inversion leads to quantitative, not qualitative, changes in face processing. *Current Biology*, *14*, 391–396. doi:10.1016/j.cub.2004.02.028

Sekunova, A., & Barton, J. (2008). Long-range and short-range relations in the perception of the vertical position of the eyes in inverted faces. *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 1129–1135. doi:10.1037/0096-1523.34.5.1129

Sergent, J. (1984). An investigation into component and configural processes underlying face perception. *British Journal of Psychology*, *75*, 221–242. doi:10.1111/j.2044-8295.1984.tb01895.x

Sergent, J. (1986). Microgenesis of face perception. In H. D. Ellis, M. A. Jeeves, F. Newcombe, & A. M. Young (Eds.), *Aspects of face processing* (pp. 17–33). Dordrecht: Martinus Nijhoff.

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., & Villemure, J. G. (1989). Prosopagnosia in a right hemispherectomized patient. *Brain*, *112*, 975–995. doi:10.1093/brain/112.4.975

Sheperd, J., Davies, G., & Ellis, H. (1981). Studies of cue saliency. In G. Davies, H. Ellis, & J. Sheperd (Eds.), *Perceiving and remembering faces* (pp. 105–131). London: Academic Press.

Sichel, J. L., & Chandler, K. A. (1969). The color-word interference test: The effects of varied color-word combinations upon verbal response latency. *Journal of Psychology*, *72*, 219–231. doi:10.1080/00223980.1969.10543502

Sinico, M. (2003). On the foundations of experimental phenomenology. *Gestalt Theory*, *25*, 111–120.

Sjoberg, W., & Windes, J. D. (1992). Recognition times for rotated normal and "Thatcher" faces. *Perceptual and Motor Skills*, *75*, 1176–1178.

Smith, M. L., Cottrell, G. W., Gosselin, F., & Schyns, P. G. (2005). Transmitting and decoding facial expressions. *Psychological Science*, *16*, 184–189. doi:10.1111/j.0956-7976.2005.00801.x

Smith, M. L., Gosselin, F., & Schyns, P. G. (2007). From a face to its category via a few information processing states in the brain. *NeuroImage*, *37*, 974–984.

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

Spillmann, L. (1999). From elements to perception: Local and global processing in visual neurons. *Perception*, *28*, 1461–1492. doi:10.1068/p2763

Spillmann, L. (2009). Phenomenology and neurophysiological correlations: Two approaches to perception research. *Vision Research*, *49*, 1507–1521. doi:10.1016/j.visres.2009.02.022

Spillmann, L., & Ehrenstein, W. H. (1996). From neuron to Gestalt-mechanisms of visual perception. In R. Greger & U. Windhorst (Eds.), *Comprehensive human physiology* (pp. 861–893). Heidelberg: Springer.

Spillmann, L., Ransom-Hogg, A., & Oehler, R. (1987). A comparison of perceptive and receptive fields in man and monkey. *Human Neurobiology*, *6*, 51–62.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, and Computers*, *31*, 137–149. doi:10.3758/BF03207704

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643–662. doi:10.1037/h0054651

Stürzel, F., & Spillmann, L. (2002). Thatcher illusion: Dependence on angle rotation. *Perception*, *29*, 937–942.

Sun, Y.-H., Ge, L., Quinn, P. C., Wang, Z., Xiao, N. G., Pascalis, O., … Lee, K. (2012). A new "fat face" illusion. *Perception*, *41*, 117–120. doi:10.1068/p6906

Susilo, T., Crookes, K., McKone, E., & Turner, H. (2009). The composite task reveals stronger holistic processing in children than adults for child faces. *PLoS ONE*, *4*(7), e6460. doi:10.1371/journal.pone.0006460

Suzuki, S., & Cavanagh, P. (1995). Facial organization blocks access to low-level features: An object inferiority effect. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 901–913. doi:10.1037/0096-1523.21.4.901

Tanaka, J., & Gordon, I. (2011). Featural, configural and holistic perception of faces. In A. Calder, J. Haxby, G. Rhodes, & M. Johnson (Eds.), *Handbook of face perception* (pp. 177–194). New York, NY: University Oxford Press.

Tanaka, J., Kay, J. B., Grinnell, E., Stansfield, B., & Szechter, L. (1998). Face recognition in young children: When the whole is greater than the sum of its parts. *Visual Cognition*, *5*, 479–496. doi:10.1080/713756795

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

Tanaka, J. W., & Farah, M. J. (2003). The holistic representation of faces. In M. A. Peterson & G. Rhodes (Eds.), *Perception of Faces, Objects, and Scenes: Analytic and Holistic Processes* (pp. 53–74). New York, NY: Oxford University Press.

Tanaka, J. W., Kaiser, M. D., Butler, S., & Le Grand, R. (2012). Mixed emotions: Holistic and analytic perception of facial expressions. *Cognition and Emotion*, *26*, 961–977. doi:10.1080/02699931.2011.630933

Tanaka, J. W., Kiefer, M., & Bukach, C. M. (2004). A holistic account of the own-race effect in face recognition: Evidence from a cross-cultural study. *Cognition*, *93*, B1–B9. doi:10.1016/j.cognition.2003.09.011

Tanaka, J. W., & Sengco, J. A. (1997). Features and their configuration in face recognition. *Memory and Cognition*, *25*, 583–592. doi:10.3758/BF03211301

Taubert, J. (2009). Chimpanzee faces are "special" to humans. *Perception*, *38*, 343–356. doi:10.1068/p6254

Taubert, J. (2010). Evidence of human-like, holistic face processing in spider monkeys (Ateles geoffroyi). *Journal of Comparative Psychology*, *124*, 57–65. doi:10.1037/a0017704

Taubert, J., & Alais, D. (2009). The composite illusion requires composite face stimuli to be biologically plausible. *Vision Research*, *49*, 1877–1885. doi:10.1016/j.visres.2009.04.025

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

Taubert, J., Apthorp, D., Aagten-Murphy, D., & Alais, D. (2011). The role of holistic processing in face perception: Evidence from the face inversion effect. *Vision Research*, *51*, 1273–1278. doi:10.1016/j.visres.2011.04.002

Taubert, J., & Parr, L. A. (2009). Visual expertise does not predict the composite effect across species: A comparison between spider (Ateles geoffroyi) and rhesus (Macaca mulatta) monkeys. *Brain and Cognition*, *71*, 187–195. doi:10.1016/j.bandc.2009.09.002

Taubert, J., Qureshi, A. A., & Parr, L. A. (2012). The composite face effect in chimpanzees (Pan troglodytes) and rhesus monkeys (Macaca mulatta). *Comparative Psychology*, *126*(4), 339–346.

Thompson, P. (1980). Margaret Thatcher: A new illusion. *Perception*, *9*, 483–484. doi:10.1068/p090483

Thompson, P. (2009). The Thatcher illusion 28 years on..... *Perception*, *38*, 921–922. doi:10.1068/pmktho

Thompson, P., & Wilson, J. (2012). Why do most faces look thinner upside down? *Iperception*, *3*, 765–774.

Thompson, S. (1880). Optical illusions of motion. *Brain*, *3*, 289–298. doi:10.1093/brain/3.3.289

Todorov, A., Loehr, V., & Oosterhof, N. N. (2010). The obligatory nature of holistic processing of faces in social judgments. *Perception*, *39*, 514–532. doi:10.1068/p6501

Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modeling of elementary psychological processes*. Cambridge: Cambridge University Press.

Treisman, A. (1996). The binding problem. *Current Opinion in Neurobiology*, *6*, 171–178. doi:10.1016/S0959-4388(96)80070-5

Turati, C., Di Giorgio, E., Bardi, L., & Simion, F. (2010). Holistic face processing in newborns, 3-month-old infants, and adults: Evidence from the composite face effect. *Child Development*, *81*, 1894–1905. doi:10.1111/j.1467-8624.2010.01520.x

Turk, M., & Pentland, A. (1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, *3*, 71–86. doi:10.1162/jocn.1991.3.1.71

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion and race in face recognition. *Quarterly Journal of Experimental Psychology*, *43A*, 161–204.

Van Belle, G., Busigny, T., Lefèvre, P., Joubert, S., Felician, O., Gentile, F., & Rossion, B. (2011). Impairment of holistic face perception following right occipito-temporal damage in prosopagnosia: Converging evidence from gaze-contingency. *Neuropsychologia*, *49*, 3145–3150. doi:10.1016/j.neuropsychologia.2011.07.010

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 Belle, G., de Graef, P., Verfaillie, K., Rossion, B., & Lefèvre, P. (2010). Face inversion impairs holistic perception: Evidence from gaze-contingent stimulation. *Journal of Vision*, *10*, 10. doi:10.1167/10.5.10

Van Wezel, R. J., & Britten, K. H. (2002). Motion adaptation in area MT. *Journal of Neurophysiology*, *88*, 3469–3476. doi:10.1152/jn.00276.2002

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & von der Heydt, R. (2012). A century of Gestalt psychology in visual perception I. Perceptual grouping and figure-ground organization. *Psychological Bulletin*, *138*, 1172–1217.

Wagemans, J., Feldman, J., Gepshtein, S., Kimchi, R., Pomerantz, J. R., van der Helm, P. A., & van Leeuwen, C. (2012). A century of Gestalt psychology in visual perception II. Conceptual and theoretical foundations. *Psychological Bulletin*, *138*, 1218–1252.

Wang, R., Li, J., Fang, H., Tian, M., & Liu, J. (2012). Individual differences in holistic processing predict face recognition ability. *Psychological Science*, *23*, 169–177. doi:10.1177/0956797611420575

Weiner, K. S., & Grill-Spector, K. (2010). Sparsely-distributed organization of face and limb activations in human ventral temporal cortex. *NeuroImage*, *52*, 1559–1573. doi:10.1016/j.neuroimage.2010.04.262

Wenger, M. J., & Ingvalson, E. M. (2002). A decisional component of holistic encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 872–892. doi:10.1037/0278-7393.28.5.872

Wenger, M. J., & Ingvalson, E. M. (2003). Preserving informational separability and violating decisional separability in facial perception and recognition. *Journal of Experimental*

Psychology: Learning, Memory, and Cognition, *29*, 1106–1118. doi:10.1037/0278-7393.29.6.1106

Wertheimer, M. (1912). Experimentelle Studien über das Sehen von Bewegung [Experimental studies on seeing motion]. *Zeitschrift für Psychologie*, *61*, 161–265.

Wertheimer, M. (1967). Gestalt theory. In W. D. Ellis (Ed.), *A source book of Gestalt psychology* (pp. 1–11). New York, NY: Humanities Press. (Original work published in 1925)

Wilmer, J. B. (2008). How to use individual differences to isolate functional organization, biology, and utility of visual functions; with illustrative proposals for stereopsis. *Spatial Vision*, *21*, 561–579. doi:10.1163/156856808786451408

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

Wong, A. C.-N., Bukach, C. M., Hsiao, J., Greenspon, E., Ahern, E., Duan, Y., & Lui, K. G. H. (2012). Holistic processing as a hallmark of perceptual expertise for nonface categories including Chinese characters. *Journal of Vision*, *12*(7), 1–15. doi:10.1167/12.7.1

Wong, A. C.-N., Bukach, C. M., Yuen, C., Yang, L., Leung, S., & Greenspon, E. (2011). Holistic processing of words modulated by reading experience. *PLoS ONE*, *6*(6), e20753. doi:10.1371/journal.pone.0020753

Wong, A. C.-N., Palmeri, T. J., & Gauthier, I. (2009). Conditions for facelike expertise with objects: Becoming a Ziggerin expert—but which type? *Psychological Science*, *20*, 1108–1117. doi:10.1111/j.1467-9280.2009.02430.x

Wong, Y. K., & Gauthier, I. (2010). Holistic processing of musical notation: Dissociating failures of selective attention in experts and novices. *Cognitive and Affective Behavioural Neuroscience*, *10*, 541–551. doi:10.3758/CABN.10.4.541

Wright, M., & Johnston, A. (1985). Invariant tuning of motion aftereffect. *Vision Research*, *25*, 1947–1955. doi:10.1016/0042-6989(85)90019-7

Wühr, P., & Waszak, F. (2003). Object-based attentional selection can modulate the Stroop effect. *Memory and Cognition*, *31*, 983–994. doi:10.3758/BF03196450

Xiao, N. G., Quinn, P. C., Ge, L., & Lee, K. (2012). Rigid facial motion influ-ences featural, but not holistic, face processing. *Vision Research*, *57*, 26–34. doi:10.1016/j.visres.2012.01.015

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, *81*, 141–145. doi:10.1037/h0027474

Yip, A., & Sinha, P. (2002). Role of color in face recognition. *Perception*, *31*, 995–1003. doi:10.1068/p3376

Young, A. W., Ellis, A. W., Flude, B. M., McWeeny, K. H., & Hay, D. C. (1986). Face-name interference. *Journal of Experimental Psychology: Human Perception and Performance*, *12*, 466–475. doi:10.1037/0096-1523.12.4.466

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

Zhou, G., Cheng, Z., Zhang, X., & Wong, A. C. (2012). Smaller holistic processing of faces associated with face drawing experience. *Psychonomic Bulletin and Review*, *19*, 157–162. doi:10.3758/s13423-011-0174-x

Zhou, H., Friedman, H. S., & von der Heydt, R. (2000). Coding of border ownership in monkey visual cortex. *Journal of Neuroscience*, *20*, 6594–6611.

Zhu, Q., Song, Y., Hu, S., Li, X., Tian, M., & Zhen, Z. (2010). Heritability of the specific cognitive ability of face perception. *Current Biology*, *20*, 137–142.

Manuscript received June 2012
Manuscript accepted January 2013
First published online May 2013