Acta Psychologica 132 (2009) 300–312

Contents lists available at ScienceDirect

Acta Psychologica

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

Distinguishing the cause and consequence of face inversion: The perceptual field hypothesis

Bruno Rossion *

Unité Cognition et Développement, Université Catholique de Louvain, 10, Place du Cardinal Mercier, 1348 Louvain-la-Neuve, Belgium

ARTICLE INFO

Article history:
Received 23 July 2009
Received in revised form 4 August 2009
Accepted 8 August 2009
Available online 10 September 2009

PsycINFO classification:
2323

Keywords:
Face inversion
Holistic perception

ABSTRACT

I published a critical review of the face inversion effect (Rossion, 2008) that triggered a few reactions and commentaries by colleagues in the field (Riesenhuber & Wolff, 2009; Yovel, in press). Here, I summarize my original paper and attempt to identify the source of both the agreements and disagreements with my colleagues, as well as other authors, regarding the nature of the face inversion effect. My view is that the major *cause* of the detrimental effect of inversion on an observer's performance at individual face recognition is the disruption of a perceptual *process*. This perceptual process is makes and observer see the multiple features of a whole individual upright face at once. It also makes the percept of a given facial feature highly dependent on the location and identity of the other features in the whole face. The perceptual process is holistic because it is driven by a holistic face *representation*, derived from visual experience. Hence, an inverted face cannot be perceived holistically: the *perceptual field* of the observer is constricted for inverted faces, each facial feature having to be processed sequentially, independently, i.e. over a smaller spatial window than the whole face. Consequently, it is particularly difficult to perceive diagnostic cues that involve several elements over a wide space on an inverted face, such as long-range relative distances between features (e.g., relative distance between eyes and mouth), or diagnostic cues that are located far away from usual gaze fixation (e.g., mouth–nose distance or mouth shape when fixating between the eyes). These difficulties are mere *consequences* of face inversion – the cause being a loss of holistic perception-, and it does not follow that relative distances between internal features are necessarily particularly important to recognize faces, that they should be labeled "configural", or should be given a specific status at the representational level. I argue that distinguishing the *cause* and *consequence(s)* of face inversion this way can provide a parsimonious and yet complete theoretical account of the face inversion effect.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

My original paper published in this journal, and which is the subject of two commentaries, was on face inversion (Rossion, 2008). An upright and an inverted face are strictly identical except for the orientation (i.e., phase) of the visual stimulus. However, we are quite good at recognizing upright faces, but terrible when the same faces are presented upside-down (Fig. 1A). This decrease of performance is known for a long time (e.g., Goldstein, 1965; Hochberg & Galper, 1967) and concerns the recognition of famous, personally familiar, or previously seen (old/new discrimination) faces. We are also significantly impaired and slowed down at matching/discriminating inverted as compared to upright unfamiliar faces (Fig. 1B), and most studies of this phenomenon have used such tasks with unfamiliar faces.

Like many other authors, I believe that a full understanding of this phenomenon will help us greatly in clarifying how the human brain perceives and recognizes faces. However, exactly 40 years after Yin's (1969) seminal study showing a much larger effect of inversion for faces than objects, the reason why face recognition is affected so much by inversion remains unclear.

I wrote this critical review of the face inversion effect mainly because I had the feeling that a series of papers published over the last few years by some of my colleagues (Riesenhuber, Jarudi, Gilad, & Sinha, 2004; Sekuler, Gaspar, Gold, & Bennett, 2004; Yovel & Kanwisher, 2004) were even increasing the conceptual confusion that currently reigns in the literature about the nature of the face inversion effect. I have now read with interest the replies of my colleagues Yovel (in press), as well as Riesenhuber and Wolff (2009) to the critical points that I raised about their original studies (Rossion, 2008). I thank these authors for their replies. In light of the complementary information that they provide in these replies, I may have made a few references to methodological aspects or results of their previous work, or the work of other authors, that were not correct, and I apologize for these mistakes (see Appendix 1). However, these are minor and largely irrelevant points, and I stand completely by the major criticisms that I made of their

* Tel.: +32 10 47 87 88; fax: +32 10 47 37 74.
E-mail addresses: bruno.rossion@uclouvain.be, bruno.rossion@psp.ucl.ac.be

0001-6918/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.actpsy.2009.08.002

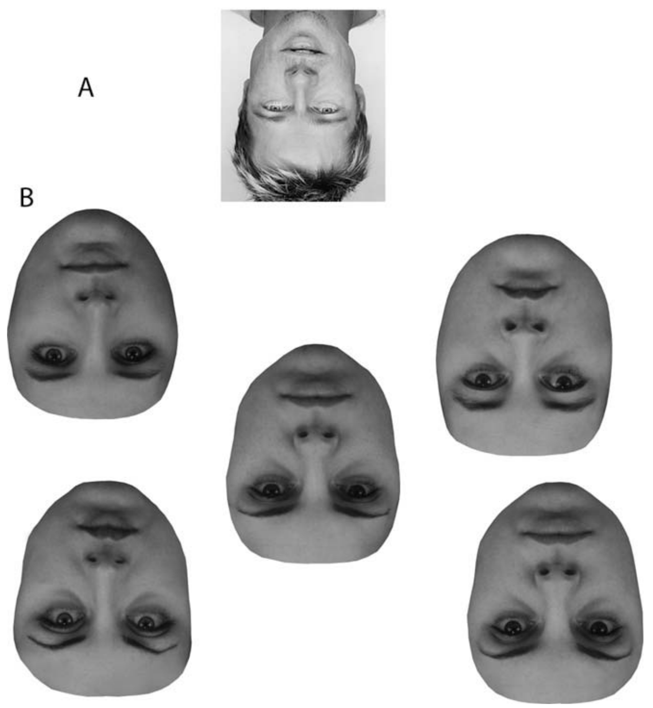

Fig. 1. (A) Inversion affects dramatically the recognition of famous faces, but also personally familiar or previously seen faces. (B) Matching or discriminating unfamiliar faces also suffers from inversion, people making more mistakes and being much slower with inverted as compared to upright faces.

original studies. Most importantly, these authors still cannot reconcile their data with the outcome of previous studies. I provide more information below, but also at the end of the present paper (Appendix 1) regarding these points, for readers interested in the precise scientific argumentation over methodological issues.

My intention for the present contribution is to be more constructive, and to aim at identifying the agreements, and most importantly the source of the remaining disagreements with these authors, in order to make progress regarding the theoretical account of the face inversion effect. I first briefly summarize what I wrote in my original paper (Rossion, 2008). I then clarify the theoretical position I took, which again, like Yovel's (in press) and Riesenhuber and Wolff's (2009) views, does not give a special status to certain diagnostic facial cues (i.e., relative distances between features) over others at the level of face representation. This is a theoretical position that is not novel at all, since it was already explicitly formulated by Tanaka and Farah (2003, pp. 62–64). However, I believe that these authors – including Tanaka and Farah – failed to consider the following: the encoding of relative distances between features could well be affected more than the encoding of local featural cues by inversion, precisely *because* integration of information over a larger spatial range is more critical to encode the former cues (relative distance) than the latter. This point, on which my colleagues and I seemingly disagree, is critical. I conclude the present paper by explaining how this view stands with respect to other theoretical positions in the field regarding the face inversion effect, and make a few suggestions for future research in this field.

2. A summary of my paper and theoretical position regarding the face inversion effect

I started my paper by claiming that, based on the literature, *what inversion does is to disrupt something called "holistic face processing*" (statement #1). This holistic processing disruption would be the major *cause* of the inversion effect (as outlined clearly previously by Tanaka, Farah and colleagues; see Farah, Drain, & Tanaka, 1995; Farah, Wilson, Drain, & Tanaka, 1998; Tanaka & Farah, 1993, 2003).

I then considered another kind of literature supporting the idea that when a face is inverted, the perception of certain cues that are diagnostic to individualize a face is more affected than the perception of other cues (statement #2). This view is what I call a qualitative view of face inversion.

In contrast, according to a *quantitative* view, all diagnostic facial cues are affected the same way by inversion. In other words, upright and inverted faces are processed just the same way, but less efficiently for inverted faces. Valentine (1988), and more recently Sekuler et al. (2004), argued in favor of this quantitative view. However, this view, other than stating that we are less good with inverted faces because we do not see faces in this orientation very often, does not offer any theoretical account of the inversion effect. I also argued in my review that showing that a local area of the face, the eyes in particular, can be the most diagnostic for processing individual faces by means of a distributed aperture method and response classification (Gosselin & Schyns, 2001; Haig, 1985), was not an argument against the view that we process faces holistically. Thus, showing that, under certain circumstances, roughly the same area of the eyes is used (Sekuler et al., 2004) and fixated (Williams & Henderson, 2007) when processing upright vs. inverted faces is an interesting observation, but it does not dismiss the qualitative view of face inversion at all.

I claimed that (#1) is the *cause* of the face inversion effect, while (#2) is merely a *consequence* of it. This cause vs. consequence relationship is a third statement made in my previous paper (#3). I consider this point as particularly important, and perhaps the only original point I made in that paper (Rossion, 2008), and which is emphasized here.

Partly because of the observation (#2), some authors have given a special status to certain face cues, at the representational level, i.e. those that are generally affected the most by inversion (e.g., Carey, 1992; Diamond & Carey, 1986; Maurer, Le Grand, & Mondloch, 2002). These authors have also *argued* that the face inversion effect was largely due to the loss of the perception of these cues, which they believed to be highly diagnostic for face recognition. In doing so, it seems that these authors confused the *cause* and the *consequence*(*s*) of the face inversion effect.

Besides an argumentation based on logic, I provided evidence from three sources supporting (#3): data from faces rotated in the plane over multiple angles; manipulation of relative distances between features in the vertical vs. horizontal direction; and the nature of the face recognition impairment in acquired prosopagnosia. The reader interested in this argumentation should refer to my previous paper (Rossion, 2008; see also Busigny & Rossion, in press; Rossion & Boremanse, 2008).

3. The source(s) of the (dis)agreements

While both Yovel (in press) and Riesenhuber and Wolff (2009) agree with me on (#1), they do not agree with me on (#2).

What we have to discuss briefly now is thus:
Whether (#2) is correct or not?
If so, why do I think that (#2) is the consequence of (#1)?

Finally, if (#2) is correct, does it imply that certain cues – in particular the relative distances between features – must have special status at the representational level?

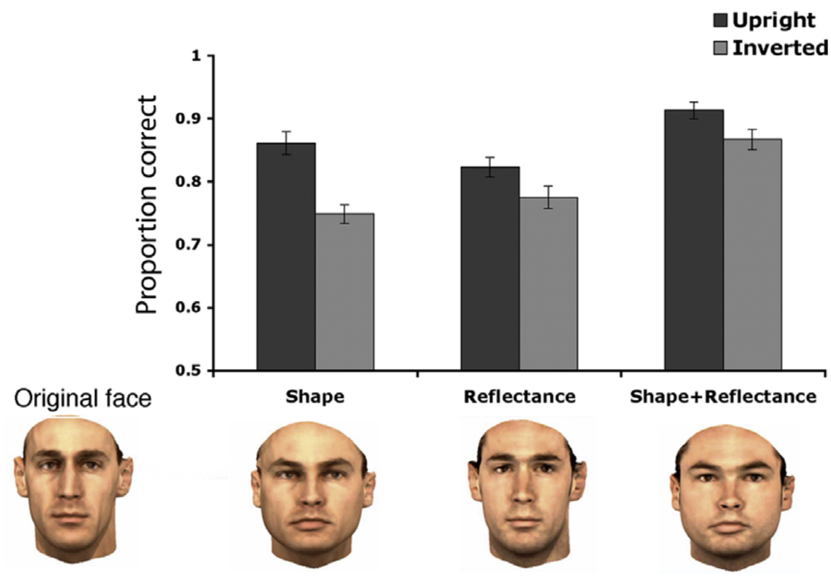

Fig. 2. Human faces vary naturally in terms of shape (defined primarily by the bone structure), and surface reflectance of the light (color and texture). Stimuli can be artificially transformed so that either shape, surface reflectance, or both can be diagnostic of identity (O'Toole et al., 1999). If natural global shape variations are preserved, the face inversion effect is larger for faces differing in shape only as compared to surface reflectance (Jiang et al., in press). This further supports the qualitative view of face inversion, according to which inversion affects certain diagnostic facial cues more than others (Rossion, 2008).

3.1. Are relative distances between features more affected – in general – by inversion than local features?

Looking at the face processing literature, it is clear to me that, *on average*, the statement (#2) is correct: when a face is inverted, certain cues that are diagnostic to individualize a face are *more* affected than other cues. What are these cues that are diagnostic to individualize a face? Individual faces vary both in terms of bone structure, defining 3D shape, and surface reflectance (brightness, color and texture) (Bruce & Young, 1998) (Fig. 2). These variations in shape and surface reflectance are either global, that is, they are found at the level of the entire face stimulus, or more local, concerning particular elements of the faces, the facial features (eyes, nose, mouth, ..., see Figs. 2 and 3). In typical 2D face stimulus as used in most experiments, that is a photograph or a schematic face, variations can be found or artificially created either by modifying the shape or the surface properties of local features (eyes elongated vs. round eyes, blue vs. brown, ...), or by modifying the metric distances between features (e.g., eyes wider apart, nose–mouth distance shorter, ...) (Fig. 3).

Among these cues that can be used quite efficiently to discriminate individual faces, which ones are more susceptible, if any, of becoming less diagnostic when a face is presented upside-down? In general, many studies have found that these diagnostic cues concern primarily the relative distances between features (e.g., nose–mouth distance; interocular distance, ...) as compared to local manipulations on these features (e.g., a modification of the shape and/or surface reflectance of the mouth or eyes for instance, without changing the feature's position). If I am correct, this larger effect of inversion for discriminating faces differing on relative distances between features than local features was observed first in a seminal study by Sergent (1984), in which she used a delayed individual face discrimination task. Since then, this effect has been shown by many independent groups of researchers, using different kinds of manipulations of features and relative distances between features (e.g., Barton, Keenan, & Bass, 2001; Boutet, Collin, & Faubert, 2003; Cabeza & Kato, 2000; Collishaw & Hole, 2000; Freire, Lee, & Symons, 2000; Goffaux, 2008; Goffaux & Rossion, 2007; Le Grand, Maurer, Mondloch, & Brent, 2001; Leder & Bruce, 2000; Leder, Candrian, Huber, & Bruce, 2001; Rhodes, Hayward, & Winkler, 2006; Searcy & Bartlett, 1996). This finding can perhaps be interpreted in different ways, but it is a fact. In short, if one considers the entire literature, both Yovel (in press) as well as Riesenhuber and Wolff (2009) are mistaken when dismissing (#2): *when a face is inverted, the perception of certain cues that are diagnostic to individualize a face is more affected than the perception of other cues.*

However, these authors (Riesenhuber et al., 2004; Yovel & Kanwisher, 2004) reported their own data, and these data do not support this claim (#2), i.e. they did not find a larger effect of inversion for manipulations of the relative distances between features than for local features in their respective experiments.

At this point, the reader should note that, in the literature, relative distances between features are often referred to as "configural" information, as opposed to "featural" information. Thus, in other words, Riesenhuber et al. (2004) found equal effects of inversion for "configural" and "featural" conditions. As a matter of fact, Yovel and Kanwisher (2004) even found a significantly larger effect of inversion for the "featural" condition than for the "configural" condition, which is the complete opposite result than above-mentioned previous studies.

What I showed in my critical review are three points: (A) the data reported in these two studies were not very convincing in my opinion, because the papers lacked important information, and there were several methodological weaknesses with the stimuli and paradigms used; (B) the methodological factors that these authors proposed to account for the discrepancy between their observations and data collected in other studies were purely speculative, and their data (i.e., Riesenhuber et al., 2004 vs. Yovel & Kanwisher, 2004) contradict each other with respect to the role of these factors; and (C) even if their results do not show a larger effect of inversion for discrimination of relative distances between features *in their experiments*, they have no good reason to rule out the rest of the literature, which largely supports (#2). I will briefly discuss these issues in turn below (see also Appendix 1, and Rossion, 2008).

(A) In their present reply, my colleagues clarify some aspects of their data, providing correct RTs for instance. It is unfortunate that the data and full analyses are still not completely available, even as supplementary material (see Appendix 1). However, it seems that, indeed, neither Yovel and Kanwisher (2004) nor Riesenhuber et al.

Fig. 3. In experiments of the face inversion effect, subtle manipulations of a base face (in the centre here) are usually made at the level of the eyes, the mouth, or both simultaneously and the participants have to discriminate these faces. These manipulations, illustrated here on a number of examples around the base face, may concern (clockwise from upper left around the base face): eyes wider apart and mouth upper in the face; eyes closer and mouth lowered; eyes lower; mouth bigger; whole eye modified (shape and surface properties); mouth color/contrast; eyeball color only; eye size. All these cues are potentially diagnostic to individualize a face, but the perception of some of these cues appear to suffer relatively more from inversion, an observation which helps us understanding the nature of the face inversion effect. Usually, an observer will fixate primarily the eye region of the face, even on an inverted face (Williams & Henderson, 2007) and will thus detect less well the variations at the level of the mouth, relatively far away from fixation, and vertical moves of the eyes, which are easily perceived by taking into account the whole face. If one fixates the mouth on these inverted faces, the relative easiness of the different discriminations change dramatically. However, on the same stimuli presented upright, the location and nature of the change matters much less, given that the face is perceived as a whole.

(2004) found a larger inversion effect for processing "configural" than "featural" trials, even when considering correct RTs. I am not going to argue against their data, but I still find some aspects of their experiments not very solid and convincing. For instance, notwithstanding the relatively small degree of these changes, that one could not assess in their original publication, Riesenhuber et al. (2004) will admit that they were not very careful in manipulating what they called the "features" of the face: they swapped entire areas around the features (eyes and mouth), which also contain diagnostic information about relative distances (a point also emphasized here by Yovel (in press)). Thus, even though the stimuli they illustrate in their reply show that these manipulations were not that large, I stand by the claim that this factor may have greatly influenced their results (see Rossion, 2008). I still suspect that this biased comparison played a role in the absence of difference between conditions, and I would have hoped that they controlled this factor in a proper experiment in their reply. Another issue is that Yovel and Kanwisher (2004) presented conditions that were not comparable because *both* the eyes and the mouth were modified in their "featural" trials, while in the majority (60%) of their "configural" trials, *either* the mouth or the eyes were modified (see Fig. 1 in Yovel and Kanwisher (2004), and Appendix 1).

Most importantly, as I indicated, Yovel and Kanwisher (2004) used only *one* male face stimulus over the course of an *entire* study. This single stimulus was transformed to give only four "featural" and four "configural" variants, and all possible combinations were used once to create pairs of stimuli to match or discriminate. Performance was thus measured in this experiment with the same faces repeated for a large amount of times, and with one kind of face shape throughout the whole experiment. The two faces in a pair differed either in terms of local features or relatives distances between these features, but face shape was normalized in that study.

Admittedly, I do not know if this normalization of the global face shape is important or not in accounting for the difference in results between Yovel and Kanwisher (2004) and other studies (see Rossion, 2008 and below). I just think that it is not ecological at all to use only a single face shape throughout an entire experiment, and it is certainly susceptible to elicit particular strategies of the participants during the task. After all, variation of the relative distances between features is naturally dependent on the global face shape (i.e., in real life circumstances, a large interocular distance is more likely to be observed in a person with a big head). In this context, it is interesting to note that variations in global face shape lead to larger inversion costs in performance than variations in internal features (Van Belle, De Smet, De Graef, Van Gool, & Verfaillie, 2009). As a matter of fact, Anaki and Bentin (submitted for publication) recently attempted to replicate Yovel and Kanwisher (2004) with a very similar design as used in the neuroimaging version of the task (blocked conditions, subjects informed about the nature of the cues) but with many more stimuli varying in shape. While Yovel and Kanwisher (2004) found a significantly larger inversion effect for "featural" than "configural" conditions, Anaki and Bentin (submitted for publication) found a non-significant trend in the other (usual) direction. This non-replication does not only cast serious doubts about the generalization of the original findings of Yovel and Kanwisher (2004), but it suggests that in this last study the lack of variation of face shape between trials may have been an important factor explaining the lack of a larger effect for "configural" trials.

(B) Neither Yovel and Kanwisher (2004) nor Riesenhuber et al. (2004) can explain yet why they found something different than other studies. In other words, they did not demonstrate that the factors that they proposed (blocking conditions, equalizing difficulty between the two kinds of trials) to explain why other studies found larger effects of inversion for "configural" trials are valid or not. In particular, Riesenhuber et al. (2004) published a paper in which they claimed that blocking or randomizing conditions led to differential costs of inversion for "configural" vs. "featural" conditions in their study. However, they had no evidence whatsoever to support this claim: they did not have a hint of a triple interaction between the factors blocking (block vs. random presentation), orientation (upright vs. inverted) and condition ("configural" vs. "featural") (see Appendix 1). Thus, they made a strong statement based on no evidence at all. I regret that Riesenhuber and Wolff (in press) do not acknowledge, in their reply, that Riesenhuber et al. (2004) did not observe any triple interaction to support their point, rather than focusing on irrelevant two-ways interactions. Yovel and Kanwisher (2004) suggested than when performance for the two conditions is equal at upright orientation, then there is no larger inversion cost for the "configural" condition. I challenged this suggestion, and I am happy to read that Yovel (in press) acknowledges that this equal performance at upright orientation is not an important, or even a contributing, factor. Thus, as I indicated previously, these factors do not appear to be crucial in explaining the discrepancy between studies. In my opinion, before making any strong claim about the factors potentially explaining the discrepancy of their results with those of previous studies, these authors should have perhaps first replicated a larger effect of inversion for "configural" vs. "featural" also with their single face stimulus, *then* show that by controlling for certain methodological factors this differential effect of inversion could be eliminated. They did not do it, and do not do it either in the present replies.

(C) As things stand, there are indeed a few studies that do not find larger effects of inversion for "configural" than "featural" modifications on faces, or studies that even find a larger effect for the latter (Yovel & Kanwisher, 2004). Yet, there is still a majority of studies that observed such a difference, some of them with a randomization of "configural" and "featural" conditions and/or equal performance at upright orientation for the two kinds of trials (see Rossion, 2008). Rather than dismissing, unfairly in my opinion, these observations, one should better pay attention to them because they are important to understand what is truly happening when an individual face is presented upside-down, i.e. the nature of the face inversion effect.

So I will conclude that, based on the literature and contrary to the claims, the statement #2 is at least partly correct: when a face is inverted, it affects the perception of certain cues more than others. *In general*, all other things being equal, it will be the relative distances between features that will be most affected. However, this detrimental effect can be potentially modulated by other factors. Yovel (in press) as well as Riesenhuber and Wolff (in press) could only speculate about these kinds of factors in their previous studies, and again they did not provide any evidence that these factors (blocking, leveling the performance for conditions at upright orientation) were critical. In subsequent papers, Yovel and Duchaine (2006) and now Yovel (in press) argue that the critical distinction to make is between studies that manipulated local shape as opposed to surface reflectance information (brightness, texture and color). I agree that this distinction is an important factor to consider, as I did in my previous paper (p. 279; see also Appendix 1 here). However, this observation is rather trivial and was made quite some time ago by researchers who manipulated surface reflectance rather than shape of the local features (e.g., Barton et al., 2001; Leder & Bruce, 2000; Searcy & Bartlett, 1996), or who, at least, minimized shape and size variations in their stimuli (e.g., Freire et al., 2000). This manipulation was done for a very good reason: local modifications of surface reflectance, contrary to local modifications of shape, do not affect the relative distances between features! Hence, we are circling around each other's reasoning: while Yovel (in press) dismisses small inversion costs for features as being due to changes in surface reflectance rather than shape, I argue that it is precisely the point: manipulations of surface reflectance do not affect relative distances. Thus, they are a more valid (i.e., independent of relative distances) manipulation of local features, at least in such experiments, than manipulations of shape and/or size.

3.1.1. The issue of shape vs. surface reflectance

How can we get around this problem? My view on this issue is the following:

(1) The human brain's efficiency in recognizing people from their face relies on multiple potential sources of diagnostic information, that is, variations between faces that characterize identity. A large amount of studies have indeed emphasized the distinction between the role of local features variations and of relative distances between these features as carrying important roles in face identity recognition. However, perhaps a more basic distinction to make is indeed between the 3D shape of the face, defined essentially but not exclusively by the bone structure of the head, and the 2D surface information (brightness, color and texture variations), defined by the reflectance of light on the skin (Bruce & Young, 1998) (Fig. 2): individual faces vary tremendously in both shape and surface reflectance information.

(2) Variations of surface reflectance are important diagnostic cues for individualizing faces (Jiang, Blanz, & O'Toole, 2006; O'Toole, Vetter, & Blanz, 1999; Russell, Biederman, Nederhouser, & Sinha, 2007). They concern the face stimulus at both the global (e.g., skin color, pigmentation), and local (eye color, mouth texture, ...) levels (Figs. 2 and 3). Hence, there is no reason to dismiss these variations as being "*primarily mediated by low-level processes*" (Yovel & Kanwisher, 2004).

(3) At a global level (i.e., the whole face), variations in surface reflectance alone lead to substantial inversion effects (Russell et al., 2007; see also Jiang, Blanz, & Rossion, in press).

(4) Even at the global level, providing that the overall contour of the head is kept diagnostic (unlike in Russell et al. (2007)) these inversion costs appear to be larger for shape than surface reflectance (Jiang et al., in press; Fig. 2). This observation supports the qualitative view of face inversion (Rossion, 2008), but it is difficult to reconcile with my colleagues' (Riesenhuber & Wolff, in press; Yovel, in press) position that all face cues suffer equally from inversion.

(5) A uniform *local* variation of surface reflectance between two face stimuli (e.g., a blue eye vs. a brown eye) is, in itself, invariant to orientation, contrary to most local variations in shape. Hence, it is understandable that variations in surface reflectance would lead to smaller inversion effects than variations of shape.

(6) Despite this invariance to orientation, local variations of surface reflectance alone *can* generally lead to small inversion costs when they are inserted in the whole face, at least in correct RTs (e.g., Barton et al., 2001; Sergent, 1984). Unfortunately, to my knowledge, the inversion effect has been measured only with variations of the eyes (color, contrast) alone or of the eyes and mouth together, but not when a feature located away from fixation such as the mouth alone for instance, would vary in color or contrast.

(7) Contrary to what is stated by Yovel (in press), even when it is only shape or size of features that varies between faces, and thus that surface reflectance is constant, inversion costs may be smaller than for modifications to relative distances between features (e.g., Malcom, Leung, & Barton, 2005).

Thus, here again, where my position differ with Yovel (in press) and Riesenhuber and Wolff (in press) and Riesenhuber et al. (2004) is that rather than dismissing these observations, I consider the fact that inversion affects differently distinct diagnostic cues to individualize faces, such as local/global shape vs. surface reflectance cues, or relative distances vs. local feature cues, as being highly interesting. It does not follow that these different cues are necessarily represented separately in the face processing system, or mediated by processes at different levels; rather it is fundamental to acknowledge and characterize their differences in orientation costs if we want to understand the nature of the face inversion effect, and more generally how we represent faces. I will turn to this issue in the next section.

3.2. Why would relative distances between features be more affected – in general – by inversion than local features?

Let me assume that #1 is correct: the loss of the ability to process holistically causes the face inversion effect. What is going to happen if you cannot process a face holistically? Well, it depends on how you define holistic processing in the first place. In my paper, I considered previous accounts, as well as my own observations, to define it as "*the simultaneous perception of the multiple features of an individual face, which are integrated into a single global representation*". There is nothing new in this definition, since, as I indicated previously, it incorporates largely the elements mentioned in earlier proposals (e.g., Farah et al., 1998; Galton, 1883; Goldstein & Chance, 1980; Ingvalson & Wenger, 2005).

This mode of processing is *functional*, in the sense that it allows an observer to encode at once all the diagnostic features of a face, including the relative distances between features, as a single representation Moreover, as Galton (1883) put it "*The general expression of a face is the sum of a multitude of small details ... If any one of them disagrees with the recollected traits of a known face, the eye is quick at observing it, and it dwells upon the difference. One small discordance overweighs a multitude of similarities and suggests a general unlikeness*." Hence, according to this view, the perception of a given facial feature depends on the whole face and vice versa. A number of experiments, which can be illustrated as visual illusions, have supported this view (e.g., Sergent, 1984; Tanaka & Farah, 1993; Thompson, 1980; Young, Hellawell, & Hay, 1987; Figs. 4 and 5). Being able to perceive a face holistically is a quite efficient way of processing a face.

It is also important to clarify a common source of confusion. "Holistic processing" here refers to a *perceptual* process: the face is *perceived* holistically, as shown in particular by the visual illusion of composite faces (Fig. 4; see also Fig. 5). Moreover, the source of behavioral effects reflecting holistic processing are found in the visual cortex, in visual areas of the right hemisphere responding preferentially to faces in particular (Harris & Aguirre, 2008; Schiltz & Rossion, 2006), and these effects arise early on during the time-course of face processing (N170 event-related potential, Bentin, McCarthy, Perez, Puce, & Allison, 1996), as early as individual face-sensitivity is observed in the human brain (Jacques, d'Arripe, & Rossion, 2007; Jacques & Rossion, 2009). Hence, faces are both *perceived* holistically and *represented* holistically. Or, to be more accurate, a face is perceived holistically precisely because the human observer has to rely on an internal face representation – which is holistic – to derive the full percept of a face.

If the face presented to the observer is upside-down, it would not be perceived holistically, as acknowledged by most authors, including Yovel (in press), and Riesenhuber and Wolff (in press). Let us consider this issue more carefully: suddenly, because the face is turned upside-down, the observer is unable to perceive the multiple features of that entire face at once, in a single global representation. That is, the observer has to analyze the features of the face one by one, independently, i.e., analytically. In other words, the observer's *perceptual field*, as I proposed previously (Rossion, 2008) becomes smaller (Fig. 6). The term "perceptual field" would refer here to *the area of vision where the observer can extract diagnostic visual information for the task*, and related terms could be the *functional visual field* or the *perceptual spatial window*.

Now, the question is the following: is this disruption of holistic processing going to affect the perception of all diagnostic facial cues equally strongly? Which diagnostic cues would be most susceptible to be affected when one has to resort on this analytical mode of face processing over a smaller perceptual field?

It seems to me that it will be the cues that necessarily involve *several* features, over a *large* space of the face, which will be affected the most (i.e., lose their diagnosticity) by inversion. For instance, comparing the distance between the mouth and the nose of two faces requires to consider the two elements altogether, over a larger space than comparing a change of color of the mouth between the two faces. Hence, it is reasonable to consider that in general discriminating two faces that differ only with respect to the nose–mouth distance will be more affected by inversion than if they differ by their mouth color, or mouth shape.

Moreover, since the perceptual field is constricted when processing an inverted stimulus, the diagnostic cues that are further away from the observer's gaze become less diagnostic. Hence, while gaze fixating on the eyes of a face does not prevent an observer to detect manipulations at the level of the mouth on an upright face, fixating the same spot on an inverted face will dramatically impair the observer's ability to detect diagnostic cues at the level of the mouth. Even though eye movements were not recorded simultaneously in these tasks, the outcome of several elegant behavioral experiments led by Barton and his colleagues support

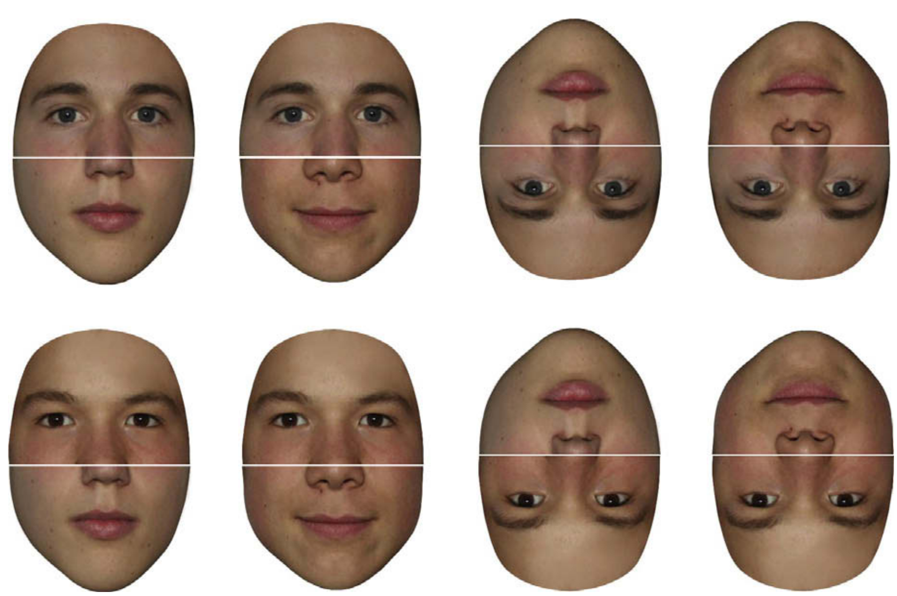

Fig. 4. The composite face illusion and inversion. When identical top halves of upright faces (cut in the middle of the nose here) are aligned with different bottom halves, one cannot help perceiving these top halves as being slightly different. This robust visual illusion indicates that the two halves of the face cannot be perceived in isolation: our perceptual system fuses the two in a single face. However, the visual illusion vanishes or is strongly attenuated when the same faces are presented upside-down, suggesting that it is our visual experience with upright faces which makes us perceiving a face as a whole. In the inverted orientation, the perceptual field being constricted, the processing of the top part of the face (appearing in the lower visual field) can be done independently of the bottom.

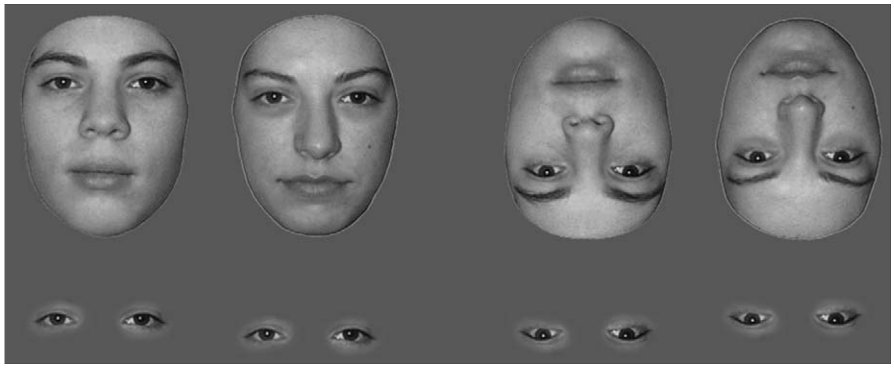

Fig. 5. Another visual illusion showing that inversion affects holistic perception of the whole face (Goffaux, 2008). When the two pairs of eyes are included in different whole faces presented at upright orientation (on the left) one would easily fail to see the eyes as being exactly the same (as shown below in isolation). However, it is much easier in inverted faces (right) because features can be perceived independently of the whole face (figure courtesy of Valérie Goffaux).

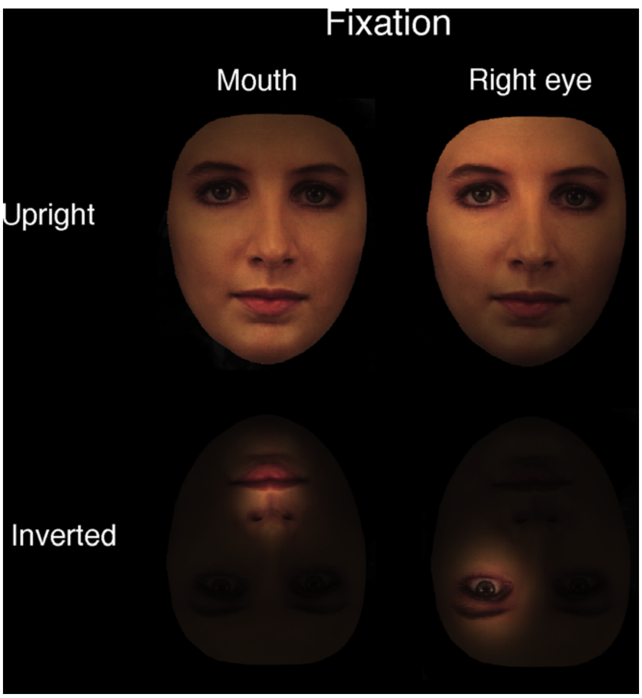

Fig. 6. Illustration of the perceptual field hypothesis in the context of holistic perception, and the face inversion effect. The observer perceives an upright face as a whole, so that different features are extracted simultaneously into a global percept, irrespective of the fixation location (*left:* mouth; *right:* right eye). However, with the same faces presented upside-down, the percept cannot be driven by a global upright face template derived from experience. Hence, the perceptual field is constricted and limited to one feature at a time (i.e., analytical processing). In general, such an analytic mode of processing will be the most detrimental for the na observer when faces differ only by long-range relative distances between features, because such diagnostic cues require to consider several elements over a wide space. When diagnostic cues are away from fixation (e.g., the mouth when fixating the eyes), they can be perceived readily in an upright face but would not fall in the perceptual field when the face is inverted. Note that the size of the perceptual field is not absolute, but relative to the size of the face (i.e., it would be smaller for a face seen from a longer distance).

this view (Barton et al., 2001; Malcom et al., 2005; Sekunova & Barton, 2008). Thus, by "other factors" playing a role in the differential inversion effect found for different types of diagnostic cues, as mentioned above, I mean that one needs to consider the relative length of the distance between features (e.g., short range vs. long-range, Sekunova & Barton, 2008) and the distance of the diagnostic cue from the observer's gaze fixation.

Thus, to me, it follows *logically*, that if a disruption of holistic perception is the primary cause (#1) of the face inversion effect, then the perception of the relative distances between features, which involve several elements of the face, over a relatively large area of the face or even over the entire face (e.g., vertical displacements of the eyes, Goffaux & Rossion, 2007; Sekunova & Barton, 2008) will suffer more than the perception of local features. In the same vein, variations in global shape of the face, including head contour, suffers more from inversion than variations in surface reflectance (Jiang et al., in press) just because the former depends more on the ability to perceive the face as a whole.

3.3. Relative distances between features depend more on holistic processing, but they do not have a special status at the representational level

If most authors agree on #1, why is it then so complicated to reach an agreement on this issue of inversion affecting differentially or not distinct facial cues?

The first reason is that a lot of other factors, alone or in combination with each other, may indeed modulate the larger effects of inversion for some cues than others. First and foremost, if relative distances between features can be handled over a relatively small spatial window (e.g., eye–eyebrow distance) they may not give rise to a large inversion effect (Sekunova & Barton, 2008). Moreover, the observer's usual gaze fixation location on the face matters a lot for an inverted face. For instance, if one fixates between the eyes of an inverted face, the constriction of the perceptual field will make it more difficult to perceive a local manipulation at the level of the mouth – far away from fixation – than a local manipulation on the eyes (Sekunova & Barton, 2008). As explained above, all these observations can be accounted for by the disruption of holistic face perception by inversion, but factors such as the observer's expectancy (and his/her fixation location) can modulate greatly the consequence of this loss of holistic perception. Hence, under certain circumstances, the perception of relative distances can be less affected by inversion than the perception of a local cue (eye–eyebrow distance vs. mouth shape to detect when observers fixate between the eyes). Yet, these circumstances and the many different factors varying between experiments are important to understand the face inversion effect, rather than ignoring these differences and claiming that equal effects of inversion should be found for all kinds of diagnostic facial cues.

As a matter of fact, methodological factors vary a lot between experiments, and will modulate the consequence of a loss of face perception. I have already discussed the factor of blocking conditions for instance, and I mentioned that in Riesenhuber et al.'s

(2004) study, blocking conditions did not interact with a differential effect of inversion for "configural" vs. "featural" trials. Nevertheless, when conditions concerning the nature (and location) of the changes to the faces are blocked, a great deal of uncertainty is removed for the observer: he/she has a better chance to be able to anticipate where and to what kind of cue to pay attention. So I agree with Riesenhuber and colleagues that blocking "configural" and "featural" in an experiment is certainly not optimal because it qualifies the disadvantage provided by the loss of holistic perception at inverted orientation. In fact, in some studies (e.g., Anaki & Bentin, submitted for publication; Yovel & Kanwisher, 2004, neuroimaging experiment), participants are *informed* about the nature and location of the changes to detect in upright and inverted orientation (e.g., "discriminate faces that differ in local parts, or relative distances between parts"). Such methodological manipulations are not ideal because, again, it reduces the advantage provided by a holistic mode of processing for upright faces. Indeed, being able to process an upright face as a whole means that *uncertainty* about the nature and location of the cues to detect is removed (see Ramon & Rossion, in press). If the experimenter removes that uncertainty by informing the participant about the nature and location of the diagnostic cues, there is a reduced cost of inversion. Similarly, many studies manipulate *both* the mouth and the two eyes *simultaneously* (either changing their shape or surface reflectance, or varying the space between them). Again, this procedure removes uncertainty, and the differential impact of holistic processing, because the observer can focus only on a small area of the face to do the task. However, when only one of these cues is changed at a time, much larger effects of inversion can be observed (e.g., Goffaux & Rossion, 2007; Malcom et al., 2005; Sekunova & Barton, 2008). Other methodological factors such as the number of faces used and their variations in head shape, or the size of the stimuli, will matter, and affect the general larger detrimental effect of face inversion for perceiving relative distances as compared to local features. It does not mean that it is impossible to understand what is primarily affected by face inversion and why, but great care should be taken when comparing studies which apparently used the same manipulations and observed different results.

The second major reason why it is difficult to reach an agreement is that some authors, Yovel in particular, appear to believe that if the diagnosticity of different facial cues is affected differentially by inversion, then it necessarily implies that these cues have a different status in the representation of faces. However, this implication is not mandatory. When a face is upright, some diagnostic cues depend on intact holistic perception more than other cues. If holistic perception is disrupted, their diagnosticity is lowered the most. This is not a good reason for one to give a specific status to these cues at the level of the holistic representation of faces. Do we give a specific status to the mouth if it is affected more by inversion – of the whole face – than the eyes (Malcom et al., 2005)? According to a strong holistic view of face processing (Tanaka & Farah, 2003) – to which I adhere by and large – the different diagnostic face cues are not stored in different subsystems. They all belong to the holistic representation of faces. In Tanaka and Farah's (2003) own words "*the holistic view maintains that featural information and configural information are not distinguishable in the face representation*" (p. 63).

In the same vein, if inversion affects relatively more the perception of changes to the mouth than the eyes (Barton et al., 2001; Malcom et al., 2005; Sekunova & Barton, 2008), should we conclude that the mouth of a face has a specific status at the representational level relative to the eyes? Should we dismiss such observations, or should we rather aim at understanding why this is the case?

Thus, I agree with my colleagues Yovel (in press) and Riesenhuber and Wolff (in press) that there is no need to make a distinction, at the level of the representation, between features and so-called "configural" (or "second-order") relations. I indicated that in the last section of my previous paper (Rossion, 2008, p. 287): "*The detrimental effect of the perception of "configural" cues following inversion is a real phenomenon, as we have seen in the main part of this paper, but it should not be taken as evidence that these cues are processed by a specific system*". I would just like to point out that this is something that was already explicitly formulated by Tanaka and Farah (1993, 2003), and thus, in due respect, that Yovel (in press) certainly cannot suggest as an original claim.

4. A summary of theoretical positions

To summarize, there are currently four theoretical positions about the nature of the face inversion effect, at least as I see it.

(1) Inversion affects all cues that are diagnostic to process faces in the same way, and this has nothing to do with holistic/configural face processing (Sekuler et al., 2004; Valentine, 1988).
(2) Inversion affects relative distances between features more than local features because these two kinds of cues do not have the same status at the representational level. Processing the former kind of cues is more important for our visual expertise with faces, and would develop particularly slowly (Carey, 1992; Carey, Diamond, & Woods, 1980; Mondloch, Le Grand, & Maurer, 2002). This loss of the ability to extract relative distances between features is one of the major causes (Barton et al., 2001; Maurer et al., 2002) or *the* cause of the inversion effect (Carey, 1992; Diamond & Carey, 1986; Leder & Bruce, 2000; Leder et al., 2001[1]).
(3) Inversion affects the perception of all diagnostic facial cues equally because they are all incorporated in the same holistic face representation (Farah et al., 1995; Riesenhuber & Wolff, in press; Riesenhuber et al., 2004; Tanaka & Farah, 1993; Yovel, in press; Yovel & Kanwisher, 2004).
(4) All diagnostic face cues are indeed incorporated in the same holistic face representation. However, in general, inversion affects the perception of relative distances between features more than local features, simply because the perception of relative distances between features depend on holistic processing relatively more than the perception of local features. Hence, the *cause* of the face inversion effect is the disruption of holistic perception, or the constriction of the perceptual field, and a *consequence* is a massive impairment of the perception of all diagnostic face cues, in particular those that involve multiple elements over a wide visual space (Rossion, 2008; the present paper).

There are many reasons why I have been arguing in favor of the fourth theoretical position. Let me explain briefly why.

The most parsimonious account is the first one (Sekuler et al., 2004; Valentine, 1988). Unfortunately, it is so parsimonious that it does not explain anything! It is also fundamentally incompatible with the most frequently reported evidence that different diagnostic cues to individualize faces are affected differentially by inver-

[1] As correctly pointed out by Riesenhuber and Wolff (in press), I acknowledge that, as many authors (e.g., Leder et al., 2001), I also subscribed to this view in a review paper of the neural correlates of the face inversion effect published a few years ago (Rossion & Gauthier, 2002). As far as I am concerned, and following a series of empirical studies on face inversion and holistic face perception carried out over the past years in my laboratory, I am now entirely convinced that this theoretical position – attributing the cause of the face inversion effect to the loss of perception of relative distances between features – is not correct and I was wrong to support it. I was quite explicit about my current theoretical view in my recent review paper about that (Rossion, 2008).

sion. This view is generally advocated by authors who consider that high-level visual perception can be independent of internal representations derived from visual experience. Thus, unsurprisingly, this view has also been associated by some authors with the idea that an inverted face has to be normalized first through mental rotation before being matched to an internal representation in memory (Collishaw & Hole, 2002; Rock, 1973; Schwaninger & Mast, 2005; Valentine, 1988; Valentine & Bruce, 1988). However, mental rotation is not necessary for face/object recognition across depth and in-plane rotations (Perrett, Oram, & Ashbridge, 1998; Wagemans, Van Gool, & Lamote, 1996), and fundamentally incompatible with empirical observations (Hayward, Zhou, Gauthier, & Harris, 2006; Rossion & Boremanse, 2008; Willems & Wagemans, 2001).

The second account (Carey, 1992; Maurer et al., 2002) is the least parsimonious because it requires making a difference of status between so-called "configural" vs. "featural" cues, at the representational level. Proponents of this view also had to distinguish between "holistic" and "configural" face processing (Carey, 1992; Maurer et al., 2002). However, there is probably no need to have "many faces of configural processing" (Carey, 1992; Maurer et al., 2002). I do not have the space to expand on this issue here, but I believe in fact that this view has led perhaps to the largest amount of confusion that we have encountered in the face processing literature over the past 15 years or so, in particular about the terms "holistic" and "configural". Originally, these two terms were used interchangeably, as synonyms of a (perceptual) process. Hence, when authors referred to "configural processing" (Sergent, 1984; Young et al., 1987) early on, they actually meant something that is defined above as "holistic processing". However, a number of influential authors have attributed a special status to relative distance between features (Carey, 1992; Carey & Diamond, 1977; Diamond & Carey, 1986; Maurer et al., 2002; Mondloch et al., 2002), which have been progressively referred to as "configural" facial cues in the literature. It is misleading in my opinion, for several reasons. First, these relative distances are not more "configural" than other aspects of the face: if a face stimulus is processed "configurally/holistically", *all* aspects of the face are "configural" (including local features and diagnostic surface reflectance cues). Second, there is no strong evidence that human observers are particularly accurate at perceiving relative distances between the fundamental internal features of the face at least (eyes, nose, mouth), especially when these variations between faces respect the variability encountered in the normal population.[2] In fact, in most studies, variations of relative distances have to be very large and unnatural (i.e., making faces grotesque) for participants to be able to reach satisfying levels of performance. We have experienced this problem in our own experiments with faces modified according to relative distances, which had to be quite large (see e.g., Fig. 1 in Goffaux and Rossion (2007)). Third, even if it observers were accurate at perceiving them, it seems that relative distances between features may not be as important to recognize an upright face than variations at the level of local features (Rhodes, 1988). That is, even if the fact that their diagnosticity being more affected by inversion is important to understand the nature of the face inversion effect, it does not mean that these relative distances between features, in particular when considering only the internal features of a face (eyes, eyebrows, nose, mouth) are that important to recognize an upright face in the first place.

In summary, this second theoretical account confounds a consequence of face inversion with its cause. It is not parsimonious at all because it suggests that they are "many faces of configural processing" (Carey, 1992; Maurer et al., 2002), while a single holistic/configural processing mode, applied to all diagnostic cues during the perception of a face may be sufficient to understand the face inversion effect.

The theoretical accounts 3 and 4 are both parsimonious: one would just need to use one term, either "holistic" or "configural" or both as synonyms, while referring to a *process* that is applied to all cues of the face. These are also theoretical positions that, contrary to the second account, do not undervalue the role of local face parts in processing faces: the local face features *are* critical of course, and their diagnosticity can only be enhanced by the other features in a holistic face representation.[3] The third account, however, is incompatible with a qualitative view of face inversion, and cannot explain why in the large majority of experiments, relative distances between features are more affected by inversion than local features. It is also incompatible with any study that reports larger effects of inversion for certain facial cues than others, e.g., even global shape as opposed to surface reflectance (brightness, color and texture) modifications (Jiang et al., in press; Fig. 2). Hence, proponents of this account are quick to dismiss surface reflectance cues as "low-level" visual information (Yovel, in press; Yovel & Kanwisher, 2004), which would not be part of the face representation. As discussed previously, this account then becomes more complex, and is certainly incompatible with a wide range of evidence indicating an important role of surface-related information in face perception (e.g., Lee & Perrett, 1997; O'Toole et al., 1999).

For all these reasons, I favor the fourth account, which appears to me as logical, parsimonious and complete: it can account for a wide range of observations related to face inversion, in particular the largest effect of inversion found for perceiving relative distances between features, but also the disruption of the composite face illusion with inversion (Figs. 4 and 5). It also accounts for the larger effects of inversion for modifications at the level of the global shape (contour) of the face as opposed to internal features (Van Belle et al., 2009) or surface reflectance (Jiang et al., in press; Fig. 2), as well as for faces which can be recognized only through the global organization of their features (faces seen from very far away, or revealed only through their lower spatial frequencies, see Boutet et al., 2003; Collishaw & Hole, 2000; Goffaux & Rossion, 2006). Moreover, it is also compatible with the fact that, as pointed out by Yovel (in press), manipulations of the shape of facial features may suffer relatively more from inversion than manipulations of local surface properties (brightness, color, texture). This is simply because the latter can be resolved locally and do not affect at all the relative distances between features. However, once again, these variations do not have to be mediated by "lower-level" visual processes (Yovel & Kanwisher, 2004).

On a final note before concluding, I do not know if this account of the face inversion effect is compatible with the simple-to-complex (i.e., hierarchical) model of face recognition proposed by Riesenhuber and colleagues (Jiang, Blanz, et al., 2006; Riesenhuber & Wolff, in press). I appreciate that these authors conceptualized the representation of faces in their model as holistic. However, the model in its present states does not seem to be able to account for the consequences of a loss of holistic face processing for inverted faces as discussed here (larger effects for relative distances,

[2] In some studies (e.g., Le Grand et al., 2001; Mondloch et al., 2002) the face stimuli are created by using the means of absolute distances between internal facial features reported in a wide anthropometric study (Farkas, 1981). Yet, it does not take into account the fact that increasing/decreasing the distance between two facial features may lead to faces that have unrealistic distances between these manipulated features, and other features of the face (Gosselin, Fortin, Michel, Schyns, & Rossion, 2007).

[3] The diagnosticity of a given feature is enhanced by the facial context (i.e., the other features in a correct face template; e.g., Tanaka & Farah, 1993), except in experimental situations where the context is not diagnostic and may interfere with the perception of the diagnostic feature (e.g., in the composite face effect, Rossion, 2008; Young et al., 1987). Both cases reflect the fact that a given feature of a face cannot be treated independently of the whole facial representation.

for distant features, or global shape variations), or the fact that this loss is abrupt (between 60° and 90° orientation) rather than gradual (Rossion & Boremanse, 2008). To me, it is not enough to acknowledge that faces are represented holistically as a result of experience (Riesenhuber & Wolff, in press). What I meant previously (Rossion, 2008; Rossion & Boremanse, 2008) is different than that. I suggested that a holistic (upright) face representation acts as a template and is necessary to guide the perception of the simple elements of an incoming visual face stimulus, in a top–down manner. Hence, the features of an upright and an inverted face would not be *perceived* the same way, because the inverted face could not benefit from this template matching. I do not see such top–down (i.e., representation-driven) processes implemented in the current hierarchical architecture proposed by Riesenhuber et al. (2004), but it would be very interesting if the model could be modified to be able to account for the full range of observations regarding the inversion effect by incorporating such processes.

5. Conclusions and suggestions for future work

In summary, the *cause* of the detrimental effect of inversion on an observer's performance at face recognition is the disruption of a process referred to as "holistic/configural" face processing. I suspect that most authors well versed in the face processing literature would agree with this view. However, to provide a more complete account of the effects observed in the literature (rather than dismissing them), I suggest that the *consequence* of face inversion is the relatively larger detrimental effect for the perception of cues that depend the most on holistic face processing: the relative distances between features, in particular when they involve multiple elements over a large space of the face. According to this view, the observer has a large perceptual or functional visual field when processing and upright face, encompassing the whole stimulus (Fig. 6). This is because the perception of the upright face stimulus is dependent on an internal representation that acts as a template for the whole face (Rossion & Boremanse, 2008). When the incoming face stimulus is upside-down, the observer cannot rely on this global internal representation, and relies on a local analysis of the features, sequentially, over a constricted perceptual field (Fig. 6).

Even though I do not expect these debates to end with the series of papers published here, I would like to express a few wishes, rather than recommendations perhaps, for future work in this area of research.

First, I wish that we could abandon the term "configural" to refer to specific cues, or sources of information on the face stimulus. This is misleading because the face is *processed* holistically/configurally, and thus *all* face cues, even local surface properties (brightness, texture, color) of an upright face are "configural" in a sense. Using the term "configural" or "configuration" only to refer to distances between features is a major source of confusion in the face processing literature. I would suggest that we rather refer explicitly to relative distances between features. The term "configural" should be better used to refer to a process, as a synonym of "holistic" (as it was used in the earlier studies, by Sergent (1984) and Young et al. (1987)).

Second, as I tried to illustrate in the last section of my previous paper (Rossion, 2008), a good way to avoid over-interpretation of data or dismiss valid observations is to know and take into consideration the history of the field of face processing, which has always balanced in cycles between a "holistic/configural" view of face perception as opposed to an emphasis on the role of local facial features (Ellis, Jeeves, Newcombe, & Young, 1986). These two views are not incompatible at all, since being able to process a face holistically enhances the diagnosticity of local facial parts. Hence, arguing that local facial parts are important for face recognition (e.g., Cabeza & Kato, 2000) does not, by any means, contradict at all the holistic perception view. The real issue will be to clarify if part-based and holistic face representations co-exist within the face processing system, perhaps in a hierarchically organized bottom-up manner as suggested by some (Jiang, Rosen, et al., 2006; Riesenhuber & Wolff, in press; Ullman, 2007), or if faces are rather first (Sergent, 1986) and only (Farah et al., 1998) processed holistically (see also Wallis, Siebeck, Swann, Blanz, and Bülthoff (2008) for a recent discussion of this issue).

Third, from a more methodological standpoint, when testing upright and inverted face processing in normal observers, I would like to reiterate the importance of measuring both accuracy and correct response times (RTs), as correct RTs could be relevant to assess differences between conditions and understand the nature of the effects observed, even when accuracy rates are not at ceiling.

Fourth, since the effect of inversion on the perception of facial cues depends on the observer's gaze fixation, we could gain critical information by systematically measuring eye fixations and saccades during upright and inverted face processing in a dynamic way (e.g., Barton, Radcliffe, Cherkasova, Edelman, & Intriligator, 2006), in particular when there are no explicit instructions about which parts of the face have to be attended.

Finally, there are a number of issues with respect to the face inversion effect that could not be addressed in this paper and in the debate with my colleagues here, but which would benefit greatly if we develop a common understanding of the nature of this phenomenon: e.g., *why is the inversion effect much larger for faces than objects?; can it be increased by visual expertise with nonface objects and under which conditions?; could it be overcome or reduced following training with inverted faces?; why it is abolished or strongly reduced following acquired prosopagnosia?; what is its developmental course?; why does it lead to increased activation in general object-sensitive brain areas of the lateral occipital cortex as well as face-sensitive event-related responses on the scalp?;* etc.

I do not know when and how these issues will be fully resolved, but I believe that one of the most important issue on the agenda of face researchers should rather be to demonstrate directly the critical aspects of holistic processing that are disrupted by inversion. Up to now, what we know from different studies is that the processing of a given feature of a face is no longer (or much less) influenced by the identity and position of the other facial features when the face is presented upside-down (e.g., Farah et al., 1998; Goffaux, 2009; Le Grand, Mondloch, Maurer, & Brent, 2004; Leder & Carbon, 2006; McKone, 2004; Rossion & Boremanse, 2008; Sergent, 1984; Suzuki & Cavanagh, 1995; Tanaka & Farah, 1993; Tanaka & Sengco, 1997; Young et al., 1987; see Figs. 4 and 5). However, a more direct demonstration of a disruption of holistic perception with inversion than this reduced interdependence of features requires the development of paradigms that could directly demonstrate the key hypothesized aspects of holistic processing and its disruption: (1) the modulation of the perceptual field (large/constricted for upright vs. inverted faces respectively); (2) the simultaneous vs. sequential perception of distinct facial features for upright and inverted faces, respectively; and (3) the integration of features into a single representation for upright but not (or much less) for inverted faces. A better understanding of how the human brain perceives, represents and recognizes faces depends on the clarification of these questions.

Acknowledgments

This work was supported by the Belgian National Fund for Scientific Research (FNRS). I would like to thank Galit Yovel, Maximilian Riesenhuber and Brian Wolff for their comments on my previous article, as well as Johan Wagemans for his comments and for giving me the opportunity to reply to my colleagues in this stimulating debate. These comments have undoubtedly helped me in clarifying my theoretical position on the issue of face inversion. I also thank Dana Kuefner, Adélaïde de Heering and Christine Schiltz for their precious comments on a previous version of this paper, Meike Ramon, Valérie Goffaux and Goedele Van Belle for their help in making the figures, and last but not least Jim Tanaka for stimulating – and ongoing – discussions on theoretical and methodological issues about face inversion and holistic perception.

Appendix 1. Specific replies to the authors' points about methods of their studies

Replies. to Yovel

1. Regarding correct RTs in the experiments of Yovel and Kanwisher (2004), it is true that I missed the supplementary material, where RT data were (partially) provided. However, I find their supplementary data to be quite confusing. For the first three experiments (behavioral performed outside the scanner, second face exemplar, and behavioral inside the scanner) the authors only report RTs for the upright, not inverted conditions, and do not report statistics. Then, they report data from an "additional study" in which they report the means (for both accuracy and RTs) only for the upright conditions of the "configural" and "featural" changes, and they report statistics for "the inversion effect" but they do not specify if those statistics came from the analyses of accuracy, RTs or a combination of both. It is unfortunate because, even with the supplementary material, one can never truly assess whether there were differential inversion costs for "configural" and "featural" trials.
2. As I indicate in the main text, contrary to what is stated by Yovel (in press), I did not overlook the difference between modifications at the level of shape of features vs. surface properties (brightness, color, texture) in my paper, which were discussed in a whole paragraph (see p. 279 in Rossion (2008)). I mentioned that these modifications were unlikely to be affected the same way by inversion than shape-changes, at least at the local level (e.g., eye color), precisely because surface reflectance modifications are largely orientation-independent, and do not change relative distances between features. However, instead of ruling out the modifications of surface properties as being '*primarily mediated by lower level processes*' (Yovel, in press; Yovel & Kanwisher, 2004, p. 895), I do not see any reason, at least for color and texture cues, to exclude them from high-level face representations. There is enough evidence that global and local surface reflectance properties play an important role in face recognition (e.g., Hill, Bruce, & Akamatsu, 1995; Lee & Perrett, 1997; O'Toole et al., 1999; Russell et al., 2007). In fact, Russell et al. (2007) found equally large effects of inversion for faces differing in terms of shape and surface reflectance for instance (when 3D shape variations were minimized, see Jiang et al., in press; and Fig. 2).
3. I do not understand how Yovel can claim that my description of their experiment (Yovel & Kanwisher, 2004) was not correct. They had a single base stimulus (#1), for which they created four "configural" faces and four "featural faces". It is true that each of these four stimuli differ according to BOTH the eyes and the mouth with respect to the target. However, in their methods section (p. 896), they indicate "*All possible combinations from the original and the four different manipulated stimuli yielded 20 different pairs for the configuration condition and 20 different pairs for the part condition. Each of the 20 different pairs of the configuration and the part sets*". This means that, if I understand correctly, for instance, for the "configural" set, we have the pairs: 1–2, 1–3, 1–4, 1–5, 2–3, 2–4, 2–5, 3–4, 3–5, 4–5 (10). In the pairs 1–2, 1–3, 1–4, 1–5, it is true that BOTH the eyes and the mouth are changing, as now described by Yovel. However, in the pair 2–3, when I look at Fig. 1 displayed by Yovel and Kanwisher (2004), only the eyes are changing horizontally, not the mouth. In other words, I was correct describing 60% of the "configural" trials used in the study of Yovel and Kanwisher (2004), and wrong for 40%. This is a serious methodological shortcoming in that study because in the "feature" condition, both the mouth and the eyes changed in each trial (100%), whereas in the "configural" condition, this proportion falls down to 40%. Moreover, for the "configural" condition, the 60% and 40% trials were all merged in the analysis, which I think is not correct because the trials in which only the mouth changes for instance, should give rise to larger inversion costs. Finally, even if I got these proportions wrong (I do not know, I can only judge based on the methods provided and the figure), I still do not see how this can be used that as an argument against the point I made in my review.

Replies. to Riesenhuber and Wolff

1. I wrote initially that "Riesenhuber et al. (2004) did not equalize performance for configural and featural trials upright". Riesenhuber and Wolff (in press) claim that I was wrong on this point, so let us look at the facts. I based this claim on their graphs and error bars, because they did not provide the statistical values in their original papers. In their main experiment, performance for "featural" trials was at 85%, while it was at 77% for "configural" trials in the experiment with random order of trials. The error bars did not overlap, but Riesenhuber and Wolff now report that this 8% difference was not significant ($p > 0.1$, with a small sample of 15 subjects) to claim that they had conditions equalized for performance at upright orientations. Based on these numbers, the reader will judge if my claim was unfounded. In any case, in their second experiment (conditions blocked), the advantage for "featural" over "configural" trials raised to about 10% ("configural first") and 21% ("featural first"). I doubt that these differences were not significant (no p-value is reported by Riesenhuber et al. (2004)), and I am surprised that the authors dare to claim that they equalized their conditions at upright orientation.
2. I mentioned in my previous paper that "there is no evidence whatsoever that this blocking factor plays any role in the absence of significantly larger inversion costs for configural than featural trials reported by Riesenhuber et al. (2004)". I stand by this point, completely. In their reply, the authors are correct to point out that in Goffaux and Rossion (2007), what was blocked was orientation, not condition. So that citation was not adequate as evidence that blocking did not affect the effect. However, what is important is precisely that "configural" and "featural" trials were randomized, as in other studies (Freire et al., 2000, experiments 3 and 4; Leder & Bruce, 2000) and that larger effects were found for "configural" trials. Most importantly, the authors still do not provide any evidence that their point was supported by their data: that is, blocking conditions did not interact at all with a differential effect of inversion for "configural" vs. "featural" trials in their study (no evidence of a triple interaction in the study of Riesenhuber et al. (2004)).
3. I thank Riesenhuber and Wolff (in press) for providing the reader with a larger set of stimuli they used. The stimuli they provide show that these manipulations were not made to the extent that I illustrated in a figure published my review. This is fine. I made this figure for illustration purposes, without making any claim about the degree of manipulations used by Riesenhuber et al. (2004). I stand by the claim that the authors

could have been more careful in making of their stimuli, because they mixed up "featural" modifications with changes of relative distances between features. At least, they should have acknowledged that it could have greatly influenced their results. I still do not know where they stand on this point, but I think that they, as well as Yovel (in press), will agree with me that when manipulating local shape of features, it affects relative distances between features as well. Even more so if one considers as a local featural modification the swapping of an entire area of the face (eyes + eyebrows) between stimuli.

4. I was also incorrect in citing Le Grand et al. (2001) as a study in which accuracy was equal at upright orientation for "configural" and "featural" trials. I referred to the original paper, in which this was the case, but there was an erratum published later (Nature 2001;412(6849):786) in which it was showed that in fact accuracy was higher on the featural set than the spacing set. I thank Cathy Mondloch for pointing this out to me.

References

Anaki, D., & Bentin, S. (submitted for publication). Time-course of featural and configural processing of faces and houses: Electrophysiological evidence.

Barton, J. J., Keenan, J. P., & Bass, T. (2001). Discrimination of spatial relations and features in faces: Effects of inversion and viewing duration. *British Journal of Psychology, 92*, 527–549.

Barton, J. J., Radcliffe, N., Cherkasova, M. V., Edelman, J., & Intriligator, J. M. (2006). Information processing during face recognition: The effects of familiarity, inversion, and morphing on scanning fixations. *Perception, 35*, 1089–1105.

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

Boutet, I., Collin, C., & Faubert, J. (2003). Configural face encoding and spatial frequency information. *Perception and Psychophysics, 65*, 1078–1093.

Bruce, V., & Young, A. W. (1998). *In the eye of the beholder: The science of face perception*. Oxford, UK: Oxford University Press.

Busigny, T., & Rossion, B. (in press). Acquired prosopagnosia abolishes the face inversion effect. *Cortex*.

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

Carey, S. (1992). Becoming a face expert. *Philosophical Transactions: Biological Sciences, 335*, 95–102.

Carey, S., & Diamond, R. (1977). From piecemeal to configurational representation of faces. *Science, 195*, 312–314.

Carey, S., Diamond, R., & Woods, B. (1980). The development of face recognition – A maturational component? *Developmental Psychology, 16*, 257–269.

Collishaw, S. M., & Hole, G. J. (2000). Featural and configurational processes in the recognition of faces of different familiarity. *Perception, 29*, 893–909.

Collishaw, S. M., & Hole, G. J. (2002). Is there a linear or a nonlinear relationship between rotation and configural processing of faces? *Perception, 31*, 287–296.

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

Ellis, H. D., Jeeves, M. A., Newcombe, F., & Young, A. W. (1986). *Aspects of face processing*. Dordrecht: Martinus Nijhof.

Farah, M., Drain, H., & Tanaka, J. (1995). What causes the face inversion effect? *Journal of Experimental Psychology: Human Perception and Performance, 21*, 628–634.

Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is "special" about face perception? *Psychological Review, 105*, 482–498.

Farkas, L. (1981). *Anthropometry of the head and face in medicine*. New York: Elsevier.

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.

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

Goffaux, V. (2008). The horizontal and vertical relations in upright faces are transmitted by different spatial frequency ranges. *Acta Psychologica, 128*, 119–126.

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

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.

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.

Goldstein, A. G. (1965). Recognition of inverted photographs of faces in children and adults. *Psychonomic Science, 3*, 447–448.

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

Gosselin, F., Fortin, I., Michel, C., Schyns, P., & Rossion, B. (2007). On the distances between internal human facial features [abstract]. *Journal of Vision, 7*(9), 499, 499a. <http://journalofvision.org/7/9/499/>, doi:10.1167/7.9.499.

Gosselin, F., & Schyns, P. G. (2001). Bubbles: A technique to reveal the use of information in recognition tasks. *Vision Research, 41*, 2261–2271.

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

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

Hayward, W. G., Zhou, G., Gauthier, I., & Harris, I. M. (2006). Dissociating viewpoint costs in mental rotation and object recognition. *Psychonomic Bull Review, 13*, 820–825.

Hill, H., Bruce, V., & Akamatsu, S. (1995). Perceiving the sex and race of faces: The role of shape and colour. *Proceedings of the Royal Society of London B, 261*, 367–373.

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

Ingvalson, E. M., & Wenger, M. J. (2005). A strong test of the dual-mode hypothesis. *Perception and Psychophysics, 67*, 14–35.

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, 9*(6), 8, 1–16. <http://journalofvision.org/9/6/8/>, doi:10.1167/9.6.8.

Jacques, C., d'Arripe, O., & Rossion, B. (2007). The time course of the inversion effect during individual face discrimination. *Journal of Vision, 7*(8), 3, 1–9. <http://journalofvision.org/7/8/3/>, doi:10.1167/7.8.3.

Jiang, F., Blanz, V., & O'Toole, A. J. (2006). Probing the visual representation of faces with adaptation: A view from the other side of the mean. *Psychological Science, 17*, 493–500.

Jiang, F., Blanz, V., & Rossion, B. (in press). Holistic processing of diagnostic 3D face shape as compared to 2D surface reflectance: Evidence from face inversion and acquired prosopagnosia. *Journal of Vision* (abstract).

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

Le Grand, R., Maurer, D., Mondloch, C. J., & Brent, H. P. (2001). Early visual experience and face processing. *Nature, 410*, 890.

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.

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

Leder, H., Candrian, G., Huber, O., & Bruce, V. (2001). Configural features in the context of upright and inverted faces. *Perception, 30*, 73–83.

Leder, H., & Carbon, C.-C. (2006). Face-specific configural processing of relational information. *British Journal of Psychology, 97*, 19–29.

Lee, K. J., & Perrett, D. (1997). Presentation-time measures of the effects of manipulation in colour space on discrimination of famous faces. *Perception, 26*, 733–752.

Malcom, G. L., Leung, C., & Barton, J. J. (2005). Regional variation in the inversion effect for faces: Differential effects for feature shape, feature configuration and external contour. *Perception, 34*, 1221–1231.

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

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.

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

O'Toole, A. J., Vetter, T., & Blanz, V. (1999). Two-dimensional reflectance and three-dimensional shape contributions to recognition of faces across viewpoint. *Vision Research, 39*, 3145–3155.

Perrett, D. I., Oram, M. W., & Ashbridge, E. (1998). Evidence accumulation in cell populations responsive to faces: An account of generalisation of recognition without mental transformations. *Cognition, 67*, 111–145.

Ramon, M., & Rossion, B. (in press). Impaired processing of relative distances between features *and* of the eye region in acquired prosopagnosia—Two sides of the same holistic coin? *Cortex*.

Rhodes, G. (1988). Looking at faces – 1st order and 2nd order features as determinants of facial appearance. *Perception, 17*, 43–63.

Rhodes, G., Hayward, W. G., & Winkler, C. (2006). Expert face coding: Configural and component coding of own-race and other-race faces. *Psychonomic Bulletin and Review, 13*, 499–505.

Riesenhuber, M., & Wolff, B. (2009). Task effects, performance levels, features, configurations, and holistic face processing: A reply to Rossion. *Acta Psychologica, 132*(3), 286–292.

Riesenhuber, M., Jarudi, I., Gilad, S., & Sinha, P. (2004). Face processing in humans is compatible with a simple shape-based model of vision. *Proceedings in Biological Sciences, 271*(Suppl. 6), S448–S450.

Rock, I. (1973). *Orientation and form*. New York: Academic Press.

Rossion, B. (2008). Picture-plane inversion leads to qualitative changes of face perception. *Acta Psychologica, 128*, 274–289.

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*, 1–13.

Rossion, B., & Gauthier, I. (2002). How does the brain process upright and inverted faces? *Behavioral and Cognitive Neuroscience Reviews, 1*, 63–75.

Russell, R., Biederman, I., Nederhouser, M., & Sinha, P. (2007). The utility of surface reflectance for the recognition of upright and inverted faces. *Vision Research, 47*, 157–165.

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

Schwaninger, A., & Mast, F. W. (2005). The face-inversion effect can be explained by the capacity limitations of an orientation normalization mechanism. *Japanese Psychological Research, 47*, 216–222.

Searcy, J. H., & Bartlett, J. C. (1996). Inversion and processing of component and spatial-relational information in faces. *Journal of Experimental Psychololögy: Human Perception and Performance, 22*, 904–915.

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.

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.

Sergent, J. (1984). An investigation into component and configural processes underlying facial perception. *British Journal of Psychology, 75*, 221–242.

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.

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.

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

Tanaka, J. W., & Farah, M. J. (2003). Holistic face recognition. In M. Peterson & G. Rhodes (Eds.), *Analytic and holistic processes in the perception of faces, objects and scenes 2* (pp. 53–91). New York: Oxford University Press.

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

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

Ullman, S. (2007). Object recognition and segmentation by a fragment-based hierarchy. *Trends in Cognitive Sciences, 11*, 58–64.

Valentine, T. (1988). Upside-down faces: A review of the effect of inversion upon face recognition. *British Journal of Psychology, 79*, 471–491.

Valentine, T., & Bruce, V. (1988). Mental rotation of faces. *Memory and Cognition, 16*, 556–566.

Van Belle, G., De Smet, M., De Graef, P., Van Gool, L., & Verfaillie, K. (2009). Configural and featural processing during face perception: A new stimulus set. *Behavior Research Methods, 41*, 279–283.

Wagemans, J., Van Gool, L., & Lamote, C. (1996). The visual system's measurement of invariants need not itself be invariant. *Psychological Science, 7*, 232–236.

Wallis, G., Siebeck, U. E., Swann, K., Blanz, V., & Bülthoff, H. H. (2008). The prototype effect revisited: Evidence for an abstract feature model of face recognition. *Journal of Vision, 8*(3), 1–15. <http://journalofvision.org/8/3/20/>, doi:10.1167/8.3.20.

Willems, B., & Wagemans, J. (2001). Matching multi-component objects from different viewpoints: Mental rotation as normalization? *Journal of Experimental Psychology: Human Perception and Performance, 27*, 1090–1115.

Williams, C.-C., & Henderson, J. M. (2007). The face inversion effect is not a consequence of aberrant eye movements. *Memory and Cognition, 35*, 1977–1985.

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

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

Yovel, G. (in press). Information about face features and the spacing among them is integrated in the holistic representation of upright faces. *Acta Psychologica* (this issue).

Yovel, G., & Duchaine, B. (2006). Specialized face perception mechanisms extract both part and spacing information: Evidence from developmental prosopagnosia. *Journal of Cognitive Neuroscience, 18*, 580–593.

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