ELSEVIER

Review

Does physical interstimulus variance account for early electrophysiological face sensitive responses in the human brain? Ten lessons on the N170

Bruno Rossion* and Corentin Jacques

University of Louvain, Belgium

A recent event-related potential (ERP) study (Thierry G., Martin, C.D., Downing, P., Pegna, A.J. 2007. Controlling for interstimulus perceptual variance abolishes N170 face selectivity. Nature Neuroscience, 10, 505–11) claimed that the larger occipito-temporal N170 response to pictures of faces than other categories – the N170 effect – is due to a methodological artifact in stimulus selection, specifically, a greater interstimulus physical variance between pictures of objects than faces in previous ERP studies which, when controlled, eliminates this N170 effect. This statement casts doubts on the validity of the conclusions reached by a whole tradition of electrophysiological experiments published over the past 15 years and questions the very interest of using the N170 to probe the time course of face processes in the human brain. Here we claim that this physical variance factor is ill-defined by Thierry et al. and cannot account for previous observations of a smaller N170 amplitude to nonface objects than faces without latency increase and component "smearing". Most importantly, this factor was controlled in previous studies that reported robust N170 effects. We demonstrate that the absence of N170 effect in the study of Thierry et al. is due to methodological flaws in the reported experiments, most notably measuring the N170 at the wrong electrode sites. Moreover, the authors attributed a modulation of N170 amplitude in their study to a differential interstimulus physical variance while it probably reflects a biased comparison of different quality sets of individual images. Here, by taking Thierry et al.'s study as an exemplar case of what should *not* be done in ERP research of visual categorization processes, we provide clarifications on a number of methodological and theoretical issues about the N170 and its largest amplitude to faces. More generally, we discuss the potential role of differential visual homogeneity of object categories as well as low-level visual properties versus high-level visual processes in accounting for early face-preferential responses and the question of the speed at which visual stimuli are categorized as faces.

This survey of the literature points to the N170 as a critical event in the time course of face processes in the human brain.

Contents

* Corresponding author. Unite Cognition et Developpement, Universite catholique de Louvain, 10, Place du Cardinal mercier, 1348 Louvain-la-Neuve, Belgium. Fax: +32 10 47 37 74.
 E-mail address: bruno.rossion@uclouvain.be (B. Rossion).
 Available online on ScienceDirect (www.sciencedirect.com).

Introduction

Recording event-related electrical potentials (ERPs) on the human scalp is a powerful and widely used method to clarify the time course of visual object categorization processes in the human

brain at the system level, non-invasively. Numerous studies have shown that pictures of faces elicit a much larger ERP of negative polarity between 130 and 200 ms than other object categories (e.g. Bentin et al., 1996; Botzel et al., 1995; Itier and Taylor, 2004a; Rossion et al., 2000; Rousselet et al., 2004a). This visual N1 component peaks at occipito-temporal electrode sites at about 170 ms following stimulus onset and has accordingly been termed the N170 (Bentin et al., 1996; Fig. 1). A similar component (M170) has been described in magnetoencephalographic studies (e.g. Halgren et al., 2000). The N170 is recorded at bilateral occipito-temporal sites and is temporally coincident with a positivity on the vertex described in earlier studies ("vertex positive potential", VPP, Botzel and Grusser, 1989; Jeffreys, 1989, 1996). The two peaks show identical response properties (Joyce and Rossion, 2005), in particular, the larger response to faces, indicating that the VPP reflects the positive counterpart of the equivalent dipoles generating the N170 in the occipito-temporal cortex (Fig. 1).

The larger response of the N170/VPP complex to faces is an undisputed observation among researchers in the field of face processing. This component has been used as a marker to investigate non-invasively the time course of face processes in the human brain in real time and characterize their sensitivity to various stimulus manipulations (*size, eccentricity*: e.g. Jeffreys et al., 1992; *isolation*
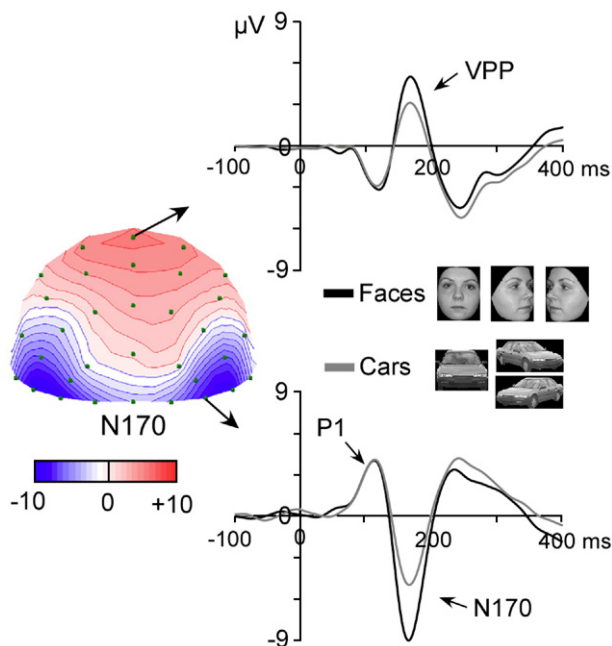

Fig. 1. The N170 is a negative component recorded from posterior lateral electrode sites following the presentation of faces and object categories (here pictures of cars). It peaks at about 160–170 ms following stimulus onset and is recorded between 130 ms and 200 ms. It is most prominent at the lowest occipito-temporal electrode sites, usually maximal on channels P8(T6) or PO8, or lower channels in this area. The component is larger in response to faces than objects in both hemispheres, with usually (but not always) a larger response in the right hemisphere. The N170 is associated with a temporally coincident positivity on the vertex (CZ), the vertex positive potential (VPP), which shows identical response properties and largely reflect the projection of the occipito-temporal dipolar sources to the vertex (see Joyce and Rossion, 2005). The data presented are grand averages of 20 subjects presented with full-front and 3/4 profiles pictures averaged together (from Rossion and Jacques, in preparation).

of features, e.g. Bentin et al., 1996; *jumbling of inner facial features*, e.g. George et al., 1996; *vertical orientation*, e.g. Rossion et al., 1999; *contrast reversal*, e.g. Itier and Taylor, 2002; *spatial frequency filtering*, e.g. Goffaux et al., 2003; *addition of visual noise*, e.g. Jemel et al., 2003) as well as task or experience factors (*perceptual adaptation*, e.g. Kovacs et al., 2006; *visual expertise*, e.g. Rossion et al., 2004; *spatial attention*, e.g. Holmes et al., 2003; *long-term familiarity*, e.g. Bentin and Deouell, 2000; *task diagnosticity*, e.g. Joyce et al., 2006; *contextual priming*, e.g. Bentin et al., 2002). However, rather surprisingly, the important and unique status that the N170 enjoys in the face literature has been recently challenged by an ERP study claiming that its largest amplitude to faces is due to an uncontrolled methodological artifact (Thierry et al., 2007a).

According to the authors of that study, previous ERP studies observed a larger N170 in response to faces (the N170 effect) because they used pictures of objects presenting a much higher interstimulus physical variance as compared to face pictures (Fig. 2A). Following the authors' reasoning, this larger physical variance would have caused an increased inter-trial jitter in the peak latency of the N170 for the nonface object conditions. Since an increase of latency jitter is known to reduce ERP amplitude (see e.g. Regan, 1989, p. 54; see also Supplementary Fig. 1), this effect would account for the smaller N170 for nonface objects. Thierry et al. further argue that when physical variance is controlled between face and object image sets, the larger N170 in response to faces is abolished (as stated by the paper's title). They provide two ERP experiments supporting this claim, which will be briefly described below.

In claiming that the N170 effect results from a mere uncontrolled artifact of inter-trial latency jitter, Thierry et al. (2007a) made a strong statement, casting doubts on the validity of the conclusions reached by tens of ERP studies published over the past 15 years. If these authors were correct, this may have important theoretical consequences in this field because it questions the very interest of studying the N170 to investigate the time course of face processes in the human brain.

Shortly after its publication, the study of Thierry et al. (2007a) was dismissed by a common brief response of several independent researchers familiar with N170 research, including the present authors (Bentin et al., 2007b). However, although the study of Thierry et al. may have failed to convince the community of N170 investigators, the claim made by these authors in a journal that has a high visibility remains an important source of confusion for a wider audience of cognitive neuroscientists, for at least two reasons. First, Thierry and colleagues seemingly replaced a complex theoretical problem ("why is the N170 larger to faces than other categories and what does this mean in terms of visual categorization processes?") with an extremely simple and easily understandable, hence appealing, solution ("because of a methodological artifact of uncontrolled physical variance between stimuli"). That is, in reality, the N170 component would not be larger in response to faces. This would be, indeed, the end of the (N170) story as far as face processing is concerned.

The second reason why Thierry et al.'s claim is confusing is implied in the title of their reply to our letter (Bentin et al., 2007b) which refers to "several intertwined perceptual and conceptual factors" (rather than the faceness of the stimulus) as accounting for the N170 amplitude (Thierry et al., 2007b). Following the original paper and this reply, a number of readers may erroneously understand Thierry et al.'s claim as being related to the debate
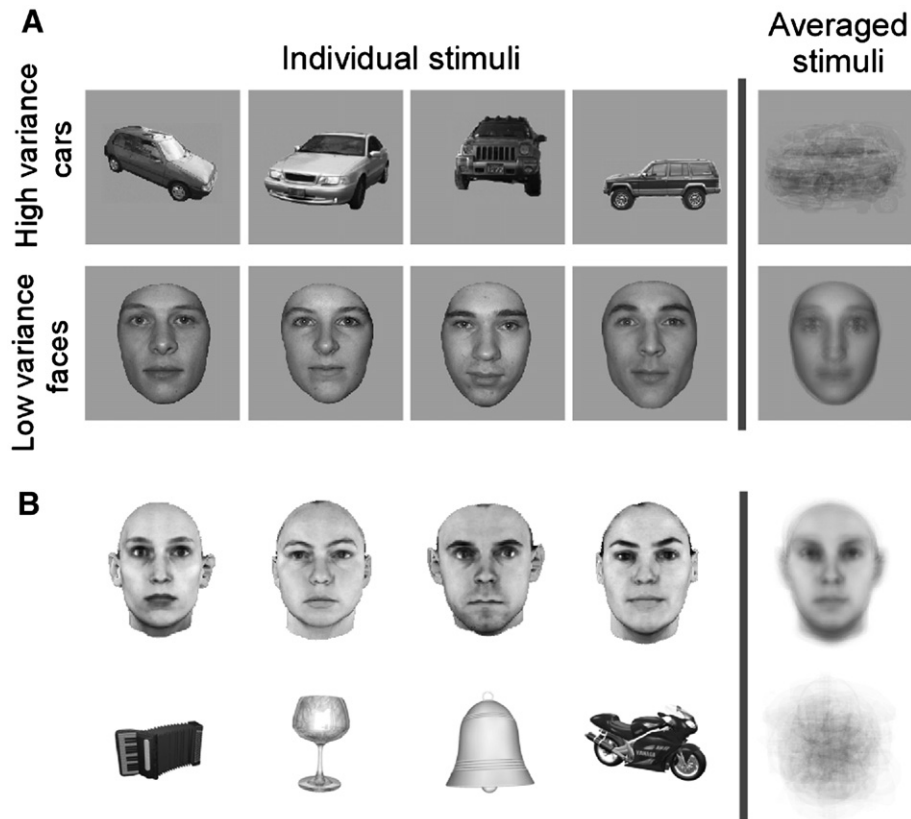
Fig. 2. (A) Thierry et al. argued that in traditional studies of the N170, while pictures of faces are presented always at the same position, viewpoint and size, there is much more physical variability in the nonface stimuli that are compared to faces (e.g. pictures of cars used in many studies). Hence, the average of 16 full-front pictures of faces can still be recognized as a face stimulus, while the average of 16 highly variable pictures of cars leads to an unrecognizable stimulus. (B) The same example with stimuli that are used in classical "face localizer" experiments in fMRI (faces vs. all kinds of object categories, e.g. Kanwisher et al., 1997), to which the criticism of Thierry et al. would be more justified (even though the activation of face-preferential responses in the visual cortex is certainly not due to a biased interstimulus physical variance factor). Only a few ERP or MEG studies used a biased comparison as displayed in B, most notably the studies inspired by the fMRI localizer approach (Liu et al., 2002, 2000), and a previous experiment by Thierry and colleagues (2006).

about the nature of the larger N170 in response to faces. This debate concerns the question of whether the processes indexed by the N170 are strictly modular (i.e., fixed) and tied to stimuli with a face geometry or if they remain flexible and can be modulated by other factors such as visual expertise for instance (see Carmel and Bentin, 2002; Rossion et al., 2002). However, importantly, debating the nature of the larger N170 to faces is a very different issue than to claim that the N170 is not larger to faces at all. Thierry et al. (2007b)'s reply appears to draw the debate in this direction but it is profoundly misleading. We believe that there is a need to clarify these issues in order to continue proper research with the N170 component, both to substantiate this debate and to increase our knowledge on the neural mechanisms and the time course of face processing in the human brain.

For these reasons, we would like to expand here on the comments that Bentin et al. (2007b) made about the study of Thierry et al. in order to shed light on the preferential N170 response to faces and its relation to physical variance of the stimulus sets. This brief rebuttal letter to Thierry et al.'s publication in *Nature Neuroscience* was entitled 'Much ado about nothing…' because, as will appear below in detail, we believe that these authors' claim is unfounded and in conflict with their own data. Moreover, because the experiments they report suffer from many theoretical and methodolo-

gical shortcomings, they do not provide any substantial evidence that could seriously question the validity of the increased N170 to faces. Yet, a potentially positive outcome of this rather unfortunate (for the community of researchers working in this area at least) publication is that it can be taken as a negative example to inform about theoretical aspects, proper methods and interpretation of N170 face experiments (i.e., a guideline from what should *not* be done in ERP research of visual categorization processes). This is the general objective of this paper, which is by no means aimed at providing an extensive review of the N170 component and does not take any strong theoretical position on its functional interpretation, only dealing with the issue of its preferential response to faces.

In ten separate sections, we will first clarify the nature of the claim made by Thierry et al. in the context of the N170 research and explain why each of these authors' claims should be dismissed in light of the existing N170 literature and the inappropriate methods they used. We will then discuss the respective weight of low-level stimulus properties and high-level visual processes in accounting for the N170 response and address the general question of the speed at which visual stimuli are categorized as faces in the human brain. Finally, we will address briefly in a last section the question that was erroneously attributed to an artifact by Thierry et al.: why is the N170 larger to faces than other categories?

Thierry et al.'s ISPV experiment

Thierry et al.'s (2007a) starting hypothesis about the larger N170 for faces than objects is that "in the great majority of previous work, it can be assumed that inter stimulus perceptual variance (ISPV) was not explicitly controlled, as no mention of this variable was made" (p. 509). These authors thus reasoned that this factor could potentially account for the larger N170 component in response to faces than objects. To test this hypothesis, they used a 2×2 factorial design with interstimulus physical variance as a factor (high vs. low) and object category as the other factor (faces vs. cars in experiment 1; faces vs. butterflies in experiment 2). They manipulated physical variance between images of a given set by presenting either full-front pictures at the same size, viewpoint and position, with a high overlap between images (low variance conditions), or all sorts of pictures differing in size, viewpoint and position (high variance conditions) (see Fig. 2A).

The N170 was larger in amplitude when resulting from an average of individual EEG responses to low variance pictures than to high variance pictures. This effect was found irrespective of the category tested (pictures of faces or cars; see Fig. 4 of Thierry et al.).

Most importantly, the N170 was larger in response to faces than cars *only* when it was confounded by the physical variance factor. That is, *low* variance *faces* gave rise to a larger N170 than *high* variance *cars*. However, according to the authors, the N170 was *not* larger for pictures of faces than cars when physical variance was controlled (high variance faces vs. cars; low variance faces vs. cars). These results were replicated in a second experiment with profile pictures of faces and butterflies compared.

Based on these results the authors concluded that the larger N170 amplitude for pictures of faces found in previous studies merely arose from averaging EEG segments elicited by highly similar pictures (faces) in which the N170 latency is highly consistent and, therefore, results in a high-amplitude average ERP. In contrast, pictures of objects being very dissimilar in previous studies (according to Thierry et al.), they led to highly inconsistent EEG segments (jittered in latency), reducing the average ERP amplitude (Fig. 2 and Supplementary Fig. 1).

What is exactly meant (and not meant) by interstimulus physical variance?

It has long been observed that, compared to other object categories, members of the face category are more similar to each other (Galton, 1883), all sharing an oval contour and internal features respecting the same basic organization (i.e., symmetrical eyes on top, above a central nose and mouth). This high visual similarity between members of the face category has for long been considered by researchers in the field as a factor to take into account when making claims about face-specific effects. For instance, some authors have considered that a greater visual similarity of individual face exemplars, coupled with the need to individualize members of the class, is a major factor accounting for specific face recognition impairments following brain damage (prosopagnosia, e.g. Damasio et al., 1982; Faust, 1955; Gauthier et al., 1999). Importantly, visual similarity among exemplars of a category is understood in these classical studies in terms of object structure, independent of variations of size, position or viewpoint of the actual pictures used in an experiment.

However, to avoid confusion, it is important to note that the claim made by Thierry et al. (2007a) in their study is different than

the question of whether there is an inherent difference in visual homogeneity between faces and objects structure that could account partly for the N170 effect. Rather, the authors' claim is that the usual N170 effect is due to the physical variance between the particular *images* used in previous experiments being much larger for objects than faces. They manipulated this variance by comparing one set of highly similar images (i.e., all full-front cars or faces centered and at similar size) to another set of highly dissimilar images (i.e., cars or faces from various viewpoints, jittered in position and of various size, see Fig. 2). In their study, Thierry et al. (2007a) quantified interstimulus physical variance by computing pixel-by-pixel correlations between the images of one condition (low or high variance). Even though there are multiple methods of assessing physical variance between images (e.g. Nederhouser et al., 2006), Thierry et al. (2007a,b) used this pixel-by-pixel correlation measure to demonstrate their point, and thus it is by considering this measure that the validity of their claim should be assessed. In summary, Thierry et al.'s claim is that the pixel-by-pixel correlation between pictures would be much greater for faces than objects in previous electrophysiological studies, leading to an increased N170 for faces. Note that this interpixel correlation between images provides a rough estimate of interstimulus physical variance, but it is intrinsically different than the *perceived* similarity of images by the visual system (i.e., requiring a psychophysical discrimination experiment). Thus, contrary to the authors' denomination, "ISPV" should not be referred to interstimulus *perceptual* variance within a class (i.e., how easily and quickly the differences among members of a visual category can be perceived by the visual system), and not even as a measure of intrinsic physical variance among members of an object class (see Fig. 2). This pixel-by-pixel correlation is just one of several possible measures of *physical* variance between a particular set of *images* used.

Is interstimulus physical variance generally uncontrolled when comparing ERP to faces and objects?

In functional magnetic resonance imaging (fMRI) studies of face processing, the so-called "face localizer" scans (Kanwisher et al., 1997) generally contrast the presentation of a set of homogenous faces to a set of objects from different categories, the latter usually having different shapes, viewpoints and surface properties (Fig. 2B). Even when stimulus are centered and size is roughly controlled, it is clear that physical image variance as defined by Thierry et al. (2007a) is usually much larger for object than face pictures in such fMRI "face localizers". Yet, one would not seriously claim that the functional brain areas lighting up as a result of this comparison in the occipito-temporal cortex, such as the 'fusiform face area' and the 'occipital face area' ('FFA', 'OFA', Kanwisher et al., 1997; see Haxby et al., 2000), reflect merely the comparison between items differing in interstimulus physical variance. Strangely enough, the criticism of Thierry et al. was rather addressed to ERP studies of face processing, despite the fact that this localizer approach is not done in traditional EEG or MEG studies. That is, EEG segments in response to objects of the *same category* are generally averaged together to obtain an average ERP response compared to the ERP in response to faces (i.e., cars vs. faces; chairs vs. faces… e.g. Bentin et al., 1996; Botzel et al., 1995; Carmel and Bentin, 2002; Eimer, 1998, 2000a,b; Goffaux et al., 2003; Itier et al., 2006; Itier and Taylor, 2004a; Jeffreys, 1996; Philiastides and Sajda, 2006; Rossion et al., 2000, 2003; Rousselet et al., 2005, 2007, 2004a; Zion-Golumbic and Bentin, 2007; MEG: Halgren et al., 2000).

Most importantly, in many of these previous ERP/MEG experiments comparing pictures of faces and other object categories, the position, size and views of pictures are constant within category, as mentioned in the methods sections of these publications (Goffaux et al., 2003; Itier et al., 2006 (see Supplementary Fig. 2 in Bentin et al., 2007b); Philiastides and Sajda, 2006; Rebai et al., 2001; Rossion et al., 2000, 2003; Rousselet et al., 2005; 2007 (see Supplementary Fig. 3 in Bentin et al., 2007b)). In addition, many other studies for which there is no reason to assume that interstimulus variance was higher for nonface objects than faces also found a larger N170 to faces (e.g. Bentin et al., 1996; Boutsen et al., 2006; Carmel and Bentin, 2002; Eimer, 1998; 2000a,b; Herrmann et al., 2005; Severac-Cauquil et al., 2000).

To illustrate this point, let us consider one of our previous studies in which we compared multiple object categories (faces, cars, shoes, greebles, chairs, houses) and found the largest N170 to faces (Rossion et al., 2000, see Fig. 3). The ERP averages were done within categories for comparison with faces, and the stimuli were carefully controlled (Fig. 3A). Within an object category, all images were presented at fixation, in the same viewpoint and at the same size. A computation of physical variance using the interstimulus pixelwise correlation measure of Thierry et al. (2007a) shows that pictures of houses had, in fact, the lowest variance among exemplars (mean interstimulus correlation close to 1), yet averaging these segments gave rise to an N170 of much lower amplitude than that for faces (Fig. 3B). Pictures of greebles and faces had identical physical variance, but the N170 was much larger in response to faces, the category of stimuli that elicited the largest N170 (Fig. 3). In another study of ours (Goffaux et al., 2003), pictures of cars and faces were undistinguishable in terms of interstimulus physical variance (Fig. 4A), yet the N170 was much larger for faces (Fig. 4B). Numerous published examples of this sort can be found in the literature (e.g. see Supplementary figures in Bentin et al., 2007b).

In short, the claim that *controlling for interstimulus perceptual variance abolishes N170 face selectivity* is not just controversial, it is incorrect: this factor was controlled for in previous studies, yet the larger N170 response to faces wa*s not* abolished at all. Although manipulating (rather than controlling, see below) this factor led to an absence of N170 amplitude difference between faces and objects in Thierry et al.'s experiments, this absence of effect does not rule out solid evidence for the largest N170 to faces collected previously by many different research groups.

In this context, it is interesting to add that performance in an individual face and individual car discrimination task with the same set of stimuli (pictures of faces and cars as used in the studies of Rossion et al., 2003; Goffaux et al., 2003 and displayed in Fig. 4A) is similarly accurate (Schiltz et al., 2006). In other words, these stimuli were in fact controlled for both *physical* variance as defined by Thierry et al. (pixel-by-pixel correlations) and *perceptual* variance (behavioral discrimination) and yet led to a much larger N170 in response to faces.

It is fair to say that measures of physical variance, as interstimulus pixelwise correlations, were not formally reported in the previous ERP studies mentioned above. However, trials from highly similar stimuli taken from the same category were averaged to obtain the N170 component, as indicated in the methods of these papers. Moreover, when the stimuli within a category are considered, there is no evidence whatsoever that factors such as position, size and viewpoint were more variable for objects than

faces.[1] Hence, a statement such as "the contrast usually reported in the literature, low ISPV faces versus high ISPV objects (Thierry et al., 2007a, p.506)" is obviously incorrect and misleading.

Regarding this issue, it is important to note that Thierry et al.'s claim is incompatible with empirically observed characteristics of the N170. For instance, if a *single* stimulus of each category (face, car, coffee mug) is used throughout the experiment so that there is no variance at all between items for any of the categories compared (Rebai et al., 2001), the N170 amplitude is still much larger for the face stimulus than for the other categories. In fact, it is well known by ERP researchers in this field that the larger N170 amplitude as compared to object categories can often be appreciated on single trials, for instance displayed as ERP images (see Rousselet et al., 2007 and section "Why is the N170 larger in amplitude to faces?"), and is thus unlikely a result of biased averaging. Finally, Thierry et al. (2007a) ignored the well-known fact that the effect of trial-to-trial jitter of latency does not only reduce the amplitude of the averaged component, but increases its latency (Regan, 1989; see Supplementary Fig. 1). Hence, if Thierry et al. (2007a) were correct, all previous studies reporting clear reduced N170 for nonface objects should also have observed delayed N170 latencies for nonface objects, which is clearly not the case (see e.g. Goffaux et al., 2003; Rossion et al., 2000; see also Fig. 5).

Considering these elements, it seems to us that in order to prove their point, Thierry et al. had to create artificial conditions and comparisons that were almost never performed in previous ERP studies of the N170. That is, they created a condition in which pictures of cars (or butterflies in their second experiment) presented at different positions, sizes and viewpoints were contrasted to faces presented in the same position, size and viewpoint (see Fig. 2A). We are not aware that a biased comparison of this order has ever been previously performed in N170 research. To our knowledge, there is only one study in which, despite averaging from within the same category, interstimulus physical variance was lower for faces than the other conditions, as acknowledged by the authors (Itier and Taylor, 2004a). In that study, a putative larger inter-trial latency jitter may indeed have further reduced the averaged N170, as suggested by the latency increase and the smearing the component for nonface objects (see Fig. 2 of Itier and Taylor, 2004a).

---

[1] While within-category physical variance is controlled in ERP studies of faces and object perception, Thierry et al. (2007b) further suggested that factors such as viewpoint or size could vary *between* categories compared and could account for the largest N170 to faces. For instance, full-front pictures of faces were sometimes compared to 3/4 views of an object category. While this is a completely different issue, there are all reasons to believe that potential differences in size and viewpoint between categories are also irrelevant in accounting for the N170 face effect. The issue of small changes of size is largely irrelevant because the N170 is mostly not affected even by large changes of size (Jeffreys, 1989, Jeffreys et al., 1992, see section "Why Thierry et al.'s manipulation of interstimulus physical variance modulated the N170 amplitude"). As for viewpoint, given that different object categories may have different canonical views (e.g. see Verfaillie and Boutsen, 1995; Blanz et al., 1999), there is no requirement or recommendation to present different object categories at the same view (e.g. all full front). As a matter of fact, the optimal viewpoint for face perception has long been defined as the 3/4 view rather than the full-front view, even though it is still a matter of debate (see Liu and Chaudhuri, 2002). In any event, several studies have compared pictures of faces to cars for instance in the same viewpoint (full-front or 3/4 profiles for the two categories) and found the larger N170 for faces (Rossion et al., 2000, 2003; Goffaux et al., 2003; see data from Rossion and Jacques, in preparation in Fig. 7).
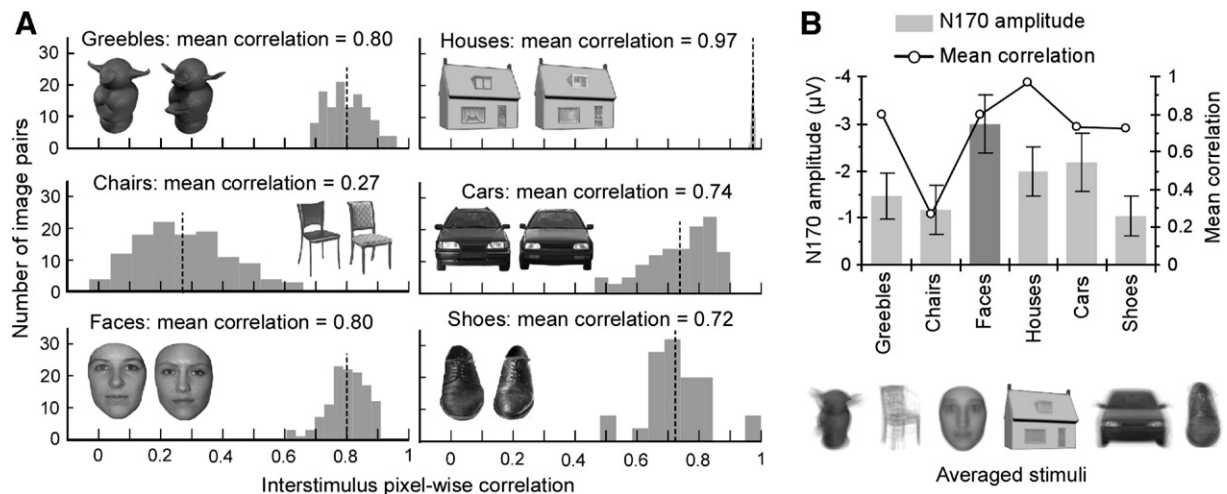
Fig. 3. (A) Distribution of pixelwise correlations (interstimulus physical variance as measured by Thierry et al.) for the 6 categories (faces, greebles, cars, houses, chairs, shoes) of stimuli used in Rossion et al. (2000). As illustrated, interpixel correlation was the highest for pictures of houses compared to the other categories ($p < 0.001$), equal between faces and greebles ($p = 0.98$), slightly lower for car and shoes which did not differ ($p = 0.14$) and lowest for chairs ($p < 0.001$). It is obvious that this factor cannot account for N170 amplitude differences, as displayed on the histogram on the right (B), in particular the larger amplitude for faces as compared to all other categories. These data were obtained from a previous study (Rossion et al., 2000) in which the N170 amplitude difference between faces and cars was only marginally significant ($p = 0.059$) with 12 participants. Here, data from 5 additional participants were included, and analyses were performed on both P8/P7 and PO8/PO7 electrodes (faces vs. cars, $p = 0.02$).

Other than that study, as indicated above, the only studies that come close to performing this kind of biased comparison are experiments in which a 'face localizer' approach (faces vs. all sorts of objects, see Fig. 2B), inspired by fMRI studies, was used. This potential methodological limitation concerned mostly MEG studies

Fig. 4. (A) In a previous ERP study manipulating spatial frequencies of visual stimuli (Goffaux et al., 2003), pictures of faces and cars were used, as in Thierry et al.'s main experiment. Contrary to the authors' claim and own stimuli sets, the stimuli of Goffaux et al. were presented under the same viewpoint and had an equal interstimulus pixelwise correlation ($p > 0.2$). Yet, as shown on B, the N170 was much larger in amplitude for pictures of faces than cars, illustrated here on a right occipito-temporal site.

(e.g. Liu et al., 2002, 2000), as well as the previous ERP study performed by Thierry and colleagues themselves with pictures of faces and objects (Thierry et al., 2006; see Fig. 1 in that paper).

Thus, contrary to Thierry et al.'s claim, there has never been a true methodological artifact with respect to interstimulus variance in the existing N170 face literature, and this claim is incompatible with existing data, such as the lack of a consistent N170 delay and smearing for nonface objects. Therefore, it appears that these authors created a straw man. As an analogy, we could claim that participants in previous ERP studies always blinked when pictures of an object category were presented, but did not blink when faces were presented. Even though there is no evidence for this, we could then ask participants of a new ERP experiment to blink in half of the trials, and consider this "blinking" as a factor in our experiment. Given that blinking will decrease the signal-to-noise ratio, we will certainly find that "the usual comparison between "face-no-blink" and "objects-blink" conditions gives rise to a larger component for faces". However, when controlling for the blinking factor (i.e., comparing face-no-blink to objects-no-blink), we would get rid of the larger potential to faces. Based on this absence of amplitude difference between faces and objects (i.e., a null effect in one study), should we then conclude against a whole literature that "controlling for blink artifacts abolishes N170 selectivity?

The N170 is larger for pictures of faces than cars

Why did Thierry et al. fail to find a significant difference between the N170 to faces and cars in their first experiment? First, an absence of a larger N170 for faces than full-front pictures of cars is not novel: it was also reported in one previous experiment (Schweinberger et al., 2004). As we pointed out previously, the N170 amplitude may vary strongly among nonface object categories (Rossion et al., 2000) and a careful look at the literature indicates that, of all nonface object categories, pictures of cars elicit the largest N170 in general (e.g. cars vs. houses and chairs in Itier et
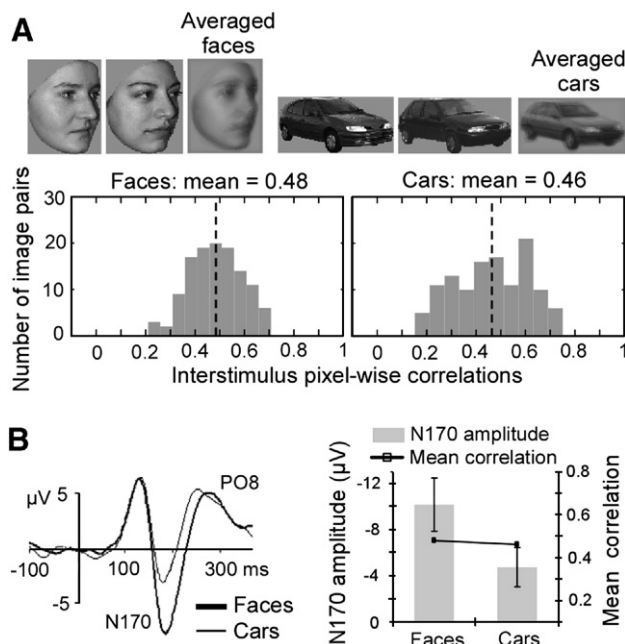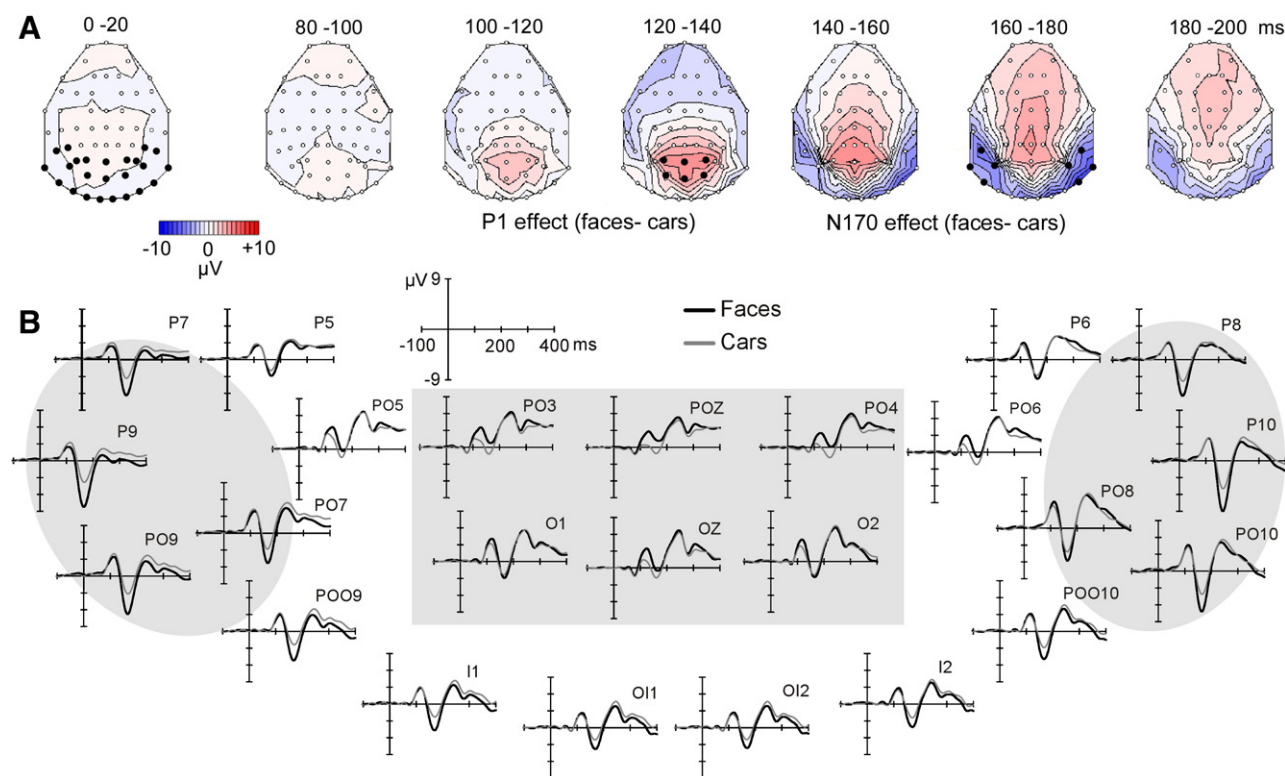
Fig. 5. (A) Topographical maps up to the first 200 ms showing the difference between electrophysiological response to faces and cars (raw data illustrated below as waveforms). All the posterior electrodes displayed on (B) are marked in black on the first topographical map on the left. There was a larger P1 in response to faces than cars, prominent between 100 and 140 ms on medial occipital and occipito-parietal sites. The six channels marked in black where this P1 effect is most prominent (120–140 ms) are highlighted in a gray rectangle below (OZ, O1, O2, PO3, PO4, POZ). They correspond to 6 out of the 10 electrodes included by Thierry et al. in their analysis of the N170, even though they do not show any N170 face effect. As illustrated both on the maps and the waveforms, they pick up most of the late P1 face effect and the posterior part of the VPP, canceling any negative amplitude difference between faces and objects (N170 face effect) observed on lateral occipito-temporal (P8/P7, PO8/PO7, P10/P9, PO10/PO9 – highlighted in gray ovals – see black marks on the topographical plots at 160–180 ms) or low occipital channels between 140 and 200 ms. Hence, by measuring the N170 on inappropriate central channels, Thierry et al. most probably cancelled out the N170 face effect. The data displayed here (Rossion and Jacques, in preparation) reflect a grand average of ERPs recorded on 20 participants (about 180 trials/condition/participant; face vs. car detection task; each stimulus repeated 6 times; common average reference). ERPs to full-front and left and right 3/4 profiles of faces and cars (from the same picture sets as in Rossion et al., 2000, 2003) were averaged together (for both faces and cars, increasing interstimulus variance equally for both categories).

al., 2006; vs. birds and furniture in Carmel and Bentin, 2002; or vs. chairs, shoes, greebles and houses in Rossion et al., 2000). Yet, due to their high familiarity, complexity and internal visual similarity comparable to faces, many studies used pictures of cars to contrast with faces; the size, position and viewpoint being kept constant within category (Bentin et al., 1996; Carmel and Bentin, 2002; Goffaux et al., 2003; Itier et al., 2006; Philiastides and Sajda, 2006; Rebai et al., 2001; Rossion et al., 2000, 2003). All these studies reported a clearly larger N170 in response to the face than car stimuli, and this includes the experiment reported by Rossion et al. (2000; see Fig. 3). This larger N170 response to face than car stimuli is also replicated in a recent study of ours (Rossion and Jacques, in preparation) in which viewpoint was included as a factor (full-front pictures, 3/4 left and right views for both faces and cars). The N170 is much larger in response to faces than cars, irrespective of the viewpoint under which the stimuli are presented (Figs. 1, 5 and 7).

It seems to us that the results obtained in two ERP studies cannot go against overwhelming evidence provided in the electrophysiological literature, especially if these two studies do not offer a reasonable account for their failure to replicate well-established previous findings. We suggest that this absence of replication is due

to the extremely small N170 amplitude reported in these two studies. In the experiment of Schweinberger et al. (2004), the N170 to pictures of human faces was less than 1.0 μV on the most sensitive channel in the right hemisphere (P8). One cannot fail to notice the parallel with the N170 measured (average of 10 electrodes) by Thierry et al. (2007a), which was of less than 0.5 μV for faces and cars in the high variance conditions, and between 1.0 and 1.5 μV in the low variance conditions (Fig. 8). In contrast to the extremely low N170 amplitudes reported by these two studies, the N170 amplitude ranges between 3.0 and 10 μV in most studies cited above (see e.g. Figs. 1, 3–7 and 9–11).

In the case of Schweinberger et al. (2004), who were rather interested in later and lower frequency components than the N170, the data were low-pass filtered with an extremely severe cutoff (10 Hz, whereas most N170 studies use a 30 to 40 Hz low-pass filter). This severe low-pass filtering certainly reduced substantially the N170 amplitude, a component that is maximal in the 5- to 15-Hz frequency band (Henson et al., 2005; Rousselet et al., 2007) (see section "Why is the N170 larger in amplitude to faces?"). Thierry et al. (2007a) did not apply a severe low-pass cutoff to their data, but other methodological shortcomings, such as averaging inadequate channels, most probably explain the small N170 and the lack of
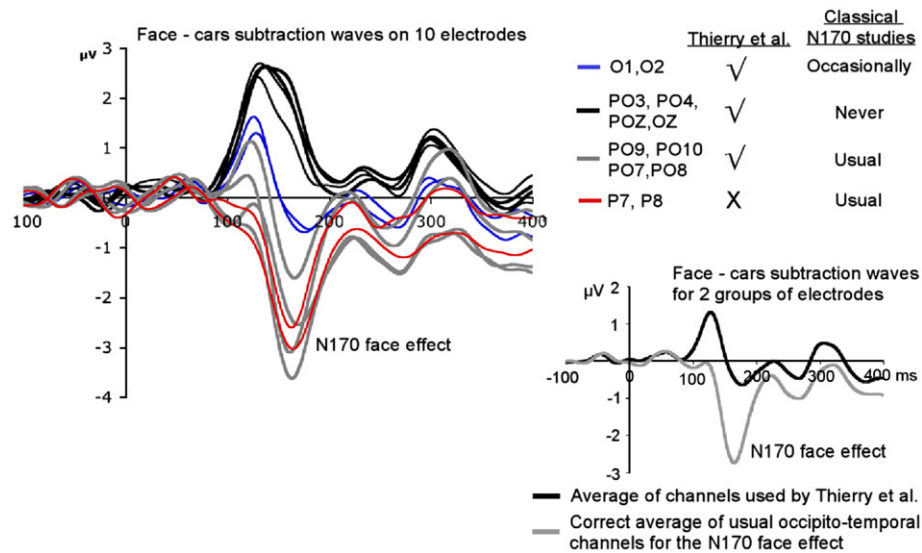
Fig. 6. *Left.* Subtraction waveforms (faces − cars) for the 10 electrode sites included by Thierry et al. in their analysis of the N170, as well as for two of their channels where the N170 effect is most sensitive but which were not included in their original analyses (P8/P7). Subtraction waveforms were computed from the data presented in Fig. 5 (Rossion and Jacques, in preparation). *Bottom right.* Averaging over the 10 channels used by Thierry et al. with our own data indeed shows a large P1 difference between faces and cars, with little or no N170 effect. In contrast, averaging differential waveforms from the correct occipito-temporal sites shows no P1 difference and a massive amplitude effect during the N170 time window.

replication of previous findings (see section "The careful and appropriate selection of channels to measure N170").

The uncontrolled interstimulus physical variance in Thierry et al.'s study

Even though the exact values of interpixel correlations for the different stimulus categories compared by Thierry et al. (2007a,b) were not reported, their figures indicate that physical variance

between faces and nonface objects was *not* controlled in their study actually, contrary to the authors' claim. Indeed, in all but one comparison in their first experiment, faces had a larger pixelwise correlation (lower physical variance) than nonface objects (Fig. 8). The absence of control is particularly conspicuous in experiment 2 as indicated by the image correlation values. That is, the mean pixelwise correlation is much higher for pictures of faces than of butterflies, and the difference between low and high physical variance conditions is much larger for faces than for butterflies
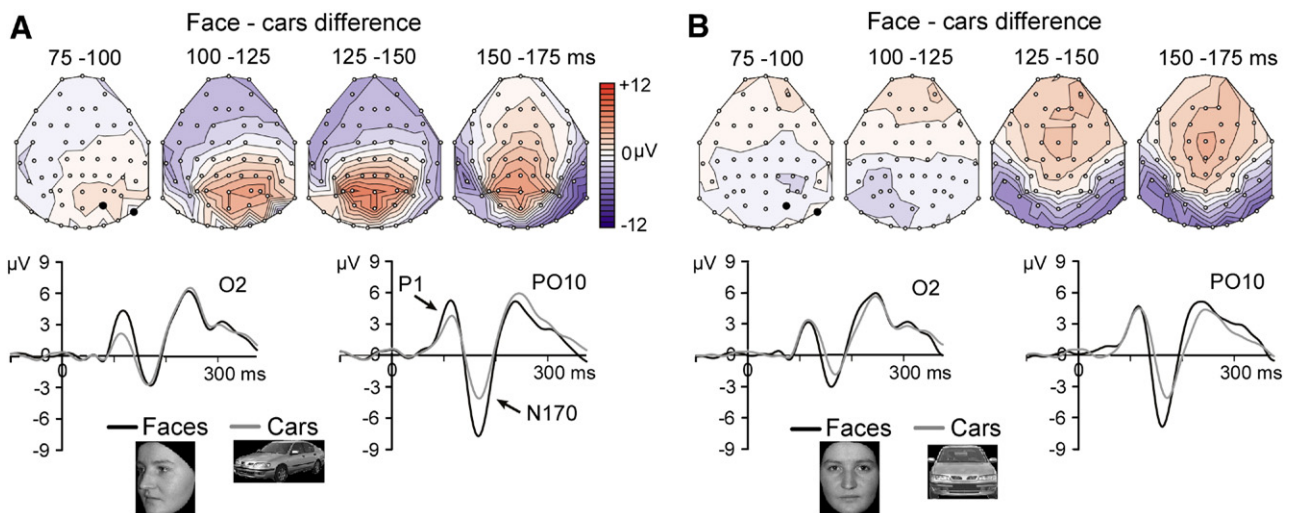

Fig. 7. (A) Grand average ERP to 3/4 view of faces and cars. Upper row shows scalp topographies of the difference between faces and cars (from 75 to 175 ms after stimulus onset). Lower row depicts raw ERPs at two posterior electrodes (location indicated in black on the left scalp topography). Note the large amplitude difference at the level of the P1 ($ps<0.001$ for the 2 channels) and N170 components ($p<0.001$ on PO10; n.s. on O2) both on ERP waveforms and scalp topographies. (B) Grand average ERP to front views of faces and cars. Details are identical to panel A. When presenting front view symmetrical stimuli, ERPs to faces and cars no longer differ at the level of the P1 component ($ps>0.3$), while the N170 is still much larger for faces. This is visible on ERP waveforms ($ps<0.005$) and scalp topographies. The fact that the N170 (but not the P1) was larger for faces irrespective of the viewpoint indicates that this effect is robust even for symmetrical full-front images of the two categories. Data are from the same experiment as depicted in Figs. 5 and 6 (Rossion and Jacques, in preparation).
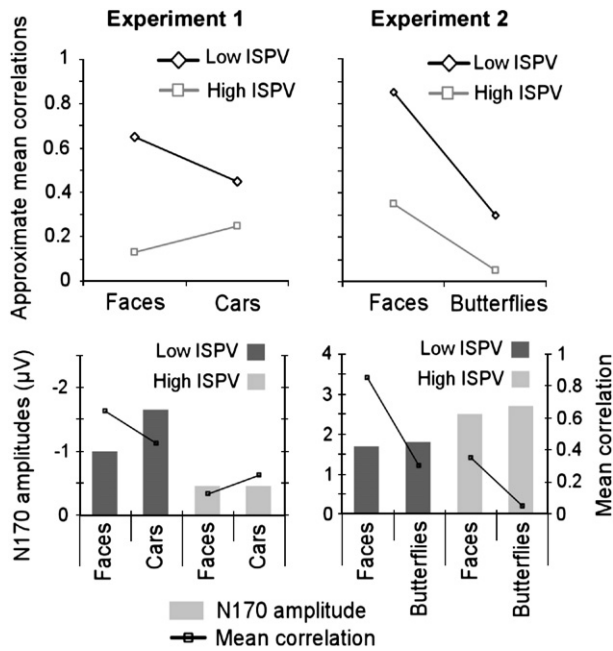
Fig. 8. While Thierry et al. claimed to have controlled for interstimulus similarity between pictures of faces and objects (exact values not reported), their own data suggest otherwise. Faces had a larger pixelwise correlation than nonface objects in all their experiments, but for high ISPV in experiment 1. The values for experiment 2 show a complete lack of control between faces and nonface objects. (B) These histograms derived from Thierry et al.'s own data provide the best illustration that ISPV cannot account for N170 amplitude, showing an inverse relationship between pixel similarity (highest for faces) and N170 amplitude (highest for cars) in the low ISPV condition (experiment 1), and no difference in N170 amplitude between categories despite a much higher pixelwise correlation for faces in the two conditions of experiment 2. Note that the amplitude values are extremely low for the N170 component (compare to Figs. 1, 3, 5, 7), for which the amplitude was averaged over 10 channels, including 4 channels (POZ, OZ, PO3, PO4) that are not sensitive at all to the N170 component itself.

(this also applies for experiment 1). In addition, histograms derived from Thierry et al.'s data show an *inverse* relationship between pixel correlation (highest for faces) and N170 amplitude (highest for cars) in the low ISPV condition (experiment 1), and no difference in N170 amplitude between categories despite a much lower variance for faces in the two conditions of experiment 2. This self-contradiction between Thierry et al.'s claim and their own values of physical variance between image sets provides the best illustration that this factor cannot account for N170 amplitude effect. Hence, whereas many previous studies did control for this interstimulus physical variance factor and reported the largest N170 to faces (see section "Is interstimulus physical variance generally uncontrolled when comparing ERP to faces and objects?"), Thierry et al. somehow managed to abolish this effect in an experiment in which the physical variance factor was in fact left completely uncontrolled.

The careful and appropriate selection of channels to measure the N170

A factor that most certainly partly accounts for the failure of Thierry et al.' study to find categorical differences on the N170 is their inadequate selection of recording sites for measuring this

component's parameters, something that is important to consider for current and future N170 investigations. Admittedly, there are no explicit guidelines in the literature as to which electrode sites should be analyzed when comparing faces to nonfaces on the N170, mainly because the current recording systems have many more electrodes than the classical 10–20 system (Jasper, 1958), and the number of electrodes recorded can vary greatly between studies (usually between 32 and 128 channels). However, it is well known that the N170 component – and its larger amplitude to faces (the N170 face effect) – has an occipito-temporal distribution on the scalp, being maximal at lateral posterior sites (e.g. P8/P7 (formerly labeled T6/T5) and PO8/PO7 (formerly CB2/CB1); or lower occipito-temporal channels when they are available: P10/P9, PO10/PO9, see Figs. 1, 5, 7). The N170 face effect falls off rapidly when recording at occipital electrode sites located close to or at the midline (O1/O2/OZ). This lateral posterior distribution is observed whether a common average reference is used, or if the reference electrode is placed on the tip of the nose, the mastoids, the earlobes or even on non-cephalic sites (see Fig. 1 in Joyce and Rossion, 2005). Consequently, researchers usually measure the N170 parameters and display their data at lateral posterior channels, where the component is maximal in amplitude (Figs. 5 and 6). Despite this well-known pattern of topographical distribution, Thierry et al. averaged data over ten electrode sites (4 in each hemisphere plus 2 midline electrodes) to measure the N170, including the medial and midline occipital electrodes (O1/O2/OZ), which were actually selected by the authors to display the N170 (absence of) effect (O1 and O2, see Figs. 4 and 6 of Thierry et al., 2007a). However, these electrodes are usually not included in N170 analyses of category effects (OZ is never included), and when analyzed or displayed the N170 effect is rarely present at these electrodes (e.g. Eimer, 2000a,b; see Figs. 5–7). Or, at least, a category x electrode interaction is reported, with the face effect being maximal at lateral sites (e.g. Itier and Taylor, 2004a, Itier et al., 2006). Unfortunately, Thierry et al. (2007a) did not report the interactions with the electrode factor.

More problematic is the fact that Thierry et al. included three additional posterior occipito-parietal electrodes (POZ, PO3, PO4) that are *never* included in conventional ERP experiments to measure the N170 face effect. Indeed, no N170 component can usually be observed on these channels (Fig. 5). In fact, rather than reflecting a negative component between 130 and 200 ms, these electrodes appear to capture most of the late part of the preceding P1 component, overlapping with the posterior part of the early VPP. That is, a positive amplitude is recorded during that time window at these channels (Figs. 5–7) rather than the negative N170! If, as in Thierry et al.'s study, the P1 (and perhaps the VPP) were larger to faces than objects (a positive polarity difference), including these electrodes in the analysis would simply cancel out any larger N170 effect for faces (a negative polarity difference) at more posterior lateral channels (PO8/PO7, PO10/PO9) also included in the analyses by Thierry et al. This cancellation of the N170 face effect has nothing to do with interstimulus variance or any other uncontrolled factor: indeed, while we observe robust differences between the N170 to faces and cars at occipito-temporal sites, this effect indeed disappears when pooling across all these inappropriate channels, similarly to what was done in Thierry et al. (Fig. 7). Hence, Thierry et al. (2007a) included 60% of N170 measures from electrodes that are not usually sensitive to the N170, going against the practice of the well-known elecrophysiological literature.

Therefore, it is not surprising that, as mentioned above, their N170 was extremely small in amplitude when averaging over these channels. In their experiment 2, the N170 was actually so small in amplitude that it had a positive polarity, something that is not reported by conventional N170 studies but which is understandable if posterior midline and occipito-parietal channels picking up the late P1 and early VPP were included in the analyses.

It is not our responsibility to clarify why Thierry et al. made such a mistake in channel selection, but one cannot help noting that in their previous and only publication on the N170 (Thierry et al., 2006), presumably with the same recording sites, they did restrict their analysis of the N170 to the correct lowest row of occipito-temporal channels (PO7/8, P7/8, PO9/10, O1/2; see layout in Fig. 5). Moreover, in line with what is customary, they displayed the N170 at the most sensitive occipito-temporal channels (PO7/8, P7/8, PO9/10), but not at O1/O2 (Thierry et al., 2006).

This bias in channel sampling in the study of Thierry et al. (2007a) was not apparent because the authors did not report any test of interactions between electrode and their independent variables (stimulus category and interstimulus variance), contrary to what is customary in ERP research (Picton et al., 2000). Furthermore, there were no topographical maps presented in the original paper and the factor hemisphere was not taken into account, again against traditional analyses, missing the important fact that the N170 amplitude difference between faces and objects is usually larger in the right than the left hemisphere (Bentin et al., 1996; Itier and Taylor, 2004a; Rossion et al., 2003; see Figs. 5, 10).

In their reply to our brief criticism of their electrode selection (Bentin et al., 2007b), Thierry and colleagues (2007b) finally displayed their waveforms from their experiment 1 at appropriate channels (P7/8, PO9/10), for all conditions separately. Interestingly, when only these low occipito-temporal electrodes (P7/8, PO9/10) are considered, the significant interaction between physical variance and stimulus category ($p < 0.01$) on all 4 electrodes appears to be largely due to the N170 being much larger for faces than cars in the high interstimulus variance conditions (see Fig. 1 of Thierry et al., 2007b)! Thus, this new analysis directly contradicts the report of the main paper of Thierry et al. (2007a) with the 10 electrodes pooled, for which these two conditions did not differ (Fig. 4c of Thierry et al., 2007a; see Fig. 8). Again, this illustrates the importance of reporting any interaction between electrode sites and the independent variable(s) of interest, in particular when emphasizing findings that depart completely from the conventional literature.

It is rather surprising that in their reply, Thierry et al. (2007b) emphasize the lack of larger N170 for faces in the low interstimulus variance conditions, an absence of difference that they attribute to their control of interstimulus variance between categories. Indeed, if anything, mean interstimulus correlation values were closer between faces and cars in their *high* interstimulus variance condition, for which there *was* a significant larger N170 to faces than cars (when the correct channels were considered; Thierry et al., 2007b) than in their low interstimulus variance condition (Fig. 8). They apparently missed this point in their reply (2007b), yet it shows that their own data contradict their original claim.

It should also be noted that the complementary segmentation analyses performed by Thierry and colleagues, which are based on a spatio-temporal cluster analysis taking into account all the electrodes recorded, do not compensate for the caveat concerning the wrong electrode selection in their analysis. First, the outcome of this segmentation approach – a method that is not widely used in

electrophysiological studies – is not purely data-driven as the authors claim but is partly dependent on the experimenter's selections of certain parameters (e.g. length of ERP time window analyzed, measure used to determine the best fit, threshold used to reject very short segments…). This bias is illustrated in the fact that the few studies that performed this analysis in the face vs. nonface ERP literature (e.g. Caldara et al., 2003; Rousselet et al., 2004b; Itier and Taylor, 2004a; Thierry et al., 2006; 2007a) reported different numbers of segment maps and differed in the number of maps explaining a single ERP component. Unfortunately, the lack of details in the description of methods and results in Thierry et al. (2007a) (e.g. topographical maps associated with each segment map, precise timing and length of each segment) prevents a possible refutation/replication of their data in an independent laboratory.

We also note that, in the few studies that used segmentation as a complementary method to ERP component analyses, face stimuli led to a qualitatively different segment map (i.e., topography) during the N170 time range (or to an 'extra' map) compared to other nonface categories (e.g. Caldara et al., 2003; Itier and Taylor, 2004a; Rousselet et al., 2004b; Thierry et al., 2006), in complete contradiction with the finding that the same segment map accounted for both faces and objects in the N170 range in Thierry et al. (2007a). Moreover, since the same N170 map was observed for faces and objects in Thierry et al. (2007a) regardless of interstimulus variance, they cannot invoke that this contradiction with previous studies comes from a difference in the control of interstimulus variance between previous studies and their most recent study (Thierry et al., 2007a). Finally, they observed the same segment during the N170 window for faces and objects and they reported an equivalent overall activation (global field power) for both categories. However, this does not mean that faces and objects were not associated with local differences (Caldara et al., 2003), which could have been masked in the global spatial correlation analyses performed on 64 channels.

In summary, as best illustrated in Figs. 6 and 8, we have all the reasons to believe that the N170 face effect was abolished by Thierry et al. because of their inappropriate selection of channels to measure the N170 rather than to a putative control of interstimulus physical variance between faces and objects.

Why Thierry et al.'s manipulation of interstimulus physical variance modulated the N170 amplitude

The N170 face effect is not due to a biased comparison of conditions that would be different with respect to interstimulus physical variance. Does it mean that this last factor is meaningless, in the sense that it plays no role at all in the differential amplitude between faces and other object categories on the N170? This remains an open question. In fact, it may well be that *part* of the N170 face effect is related to this factor. However, as we have indicated above (The uncontrolled interstimulus physical variance in Thierry et al.'s study), Thierry et al. did not test and manipulate that factor properly. As noted above, their measure of physical similarity, by computing interpixel correlations between pairs of images, is not adequate when comparing stimuli from different categories. For instance, pixelwise correlation is highly dependent on largely irrelevant factors such as the size of the background frame in which the stimulus is depicted and the contrast between this background frame and the stimulus (i.e., a large contrast with the background yields a much higher correlation due to larger

variance in pixel luminance throughout the image).[2] Another trivial example should help understand why this measure is not appropriate in this context: if a single object is used throughout an experiment but is slightly varied in size from trial to trial, physical similarity as computed by Thierry et al. will decrease dramatically, even though the very same stimulus is presented. However, as demonstrated by Jeffreys and colleagues (1992), the size of the stimuli is largely irrelevant when considering the amplitude of these high-level visual components. More precisely, an eightfold size change does not greatly affect the amplitude of the VPP (the positive counterpart of the N170) evoked by optimally fixated faces (Jeffreys, 1989; Jeffreys et al., 1992). This remains so unless the fixation point is moved off the face, leading to a substantial reduction of the component's amplitude (Jeffreys et al., 1992; for the N170, see Rousselet et al., 2005). This size invariance of the N170 is understandable since this component presumably reflects mainly high-level visual processes that occur at least 130 ms after stimulus onset. In this context, it is difficult to find a theoretical justification as to why the amplitude of this component would be highly sensitive to a factor such as the size of the image, which nevertheless increases dramatically interstimulus physical variance as computed by Thierry et al. (2007a).

There are other simple reasons why the physical variance measure used by Thierry et al. should not be recommended for future studies when comparing different stimulus categories. For instance, this measure is dependent on the complexity of the object structure, with higher correlations for faces compared to objects that have external parts such as many man-made objects or even animals. It is also highly dependent on the spatial frequency content of the stimulus, with much lower correlation between pixels of two images if these contain a large amount of information in high spatial frequency bands for instance (hence the lower mean correlation for chairs than other categories in Rossion et al., 2000; Fig. 3).

Given the inadequacy of this measure, the reason why Thierry et al. observed a significant effect of interstimulus physical variance (independent of the category) on the N170 remains unclear. In their first experiment, they contrasted conditions in which the images were visually homogenous to conditions where they varied in size, viewpoint and eccentricity (see Fig. 2A; see also Fig. 1 in Thierry et al., 2007a). According to them, the latter conditions give rise to a less consistent N170 latency across trials (increase of inter-trial jitter), leading to reduced N170 amplitude (Fig. 2). However, Thierry et al. did not provide any evidence that inter-trial jitter – as measured by phase-locking (e.g. Henson et al., 2005; Rousselet et al., 2007) – was decreased around the N170 time window for the high variance as compared to the low variance conditions. As indicated earlier, the putative effect of a larger jitter between trials evoked by high variance stimuli should have caused not only an amplitude reduction but also an increase of latency and a "smearing" of the component for this condition (see Regan, 1989). There was no evidence for such a latency delay in their experiment 1 at least and no hint of a wider N170 component in Thierry et al.'s data. Hence, while it may be that a small increase of inter-trial latency jitter occurred in the high variance conditions of that study indeed, this must have been minimal and this factor alone is unlikely to

account for much of the reduction of N170 amplitude. Other factors must be at play.

To put it simply, we believe that the most parsimonious explanation of the "physical variance effect" found for both faces and cars in Thierry et al.'s experiment is that it reflects merely a stimulus effect. That is, the authors compared one set of high quality individual images (all full-front faces and full-front cars) to another set of lower quality images (faces and cars at different sizes and viewpoint). Possibly, the N170 was larger in response to the individual images of the first set because these individual images, irrespective of their relation with each other, give rise to larger N170 responses than the individual images of the second set. This is a most plausible account but a totally uninteresting one: Thierry et al. compared two different sets of quality stimuli leading to different N170 amplitudes, and this would explain the effect that they attributed to interstimulus physical variance. In fact, most probably, the lower N170 amplitude observed for individual images in the high variance condition was found because for a substantial part of these stimuli (faces or cars), the image presented contained fewer category diagnostic information (i.e., profile views or perhaps even back views), eliciting small N170s, as individual trials.

An additional factor that may also be critical in accounting for Thierry et al.'s observations, independently of the quality of each picture used in the experiments, is the eccentricity of the diagnostic features with respect to the fixation point. A progressively slightly later but mainly smaller VPP/N170 peak is evoked for increasingly eccentric presentation of the stimulus away from fixation toward the left or right visual field (for VPP, see Jeffreys et al., 1992; for N170 see Rousselet et al., 2005; as well as Jacques and Rossion, 2004, 2007b). Because in each block of 100 trials used by Thierry et al., there were at least half of the trials with full-front stimuli presented at the center (the low variance trials), the most efficient strategy for participants to perform the one back task requested was to fixate the center of the screen (as participants were probably instructed to do anyway). Thus, individual pictures in the low variance condition fell exactly on the fovea, eliciting an N170 of maximal amplitude and an early latency in each trial. In contrast, for a large subset of the pictures used in the high variance condition, diagnostic features appeared slightly away from the fovea, leading to a smaller and slightly delayed N170 on most individual trials compared to the low interstimulus variance condition. Hence, even though the exact condition of stimulation is impossible to predict from trial to trial in Thierry et al.'s experiments, comparing trials for which the diagnostic features always fall on the fovea (low variance) to trials for which these features are most often out of the fovea (high variance) can lead to an amplitude difference on the N170, independent from and in addition to a possible small effect of jitter caused by the latency delay.

The same explanation holds for their experiment 2 where profile views of faces and butterflies were compared. There, rather than taking two different sets of images for low and high variance conditions, Thierry et al. artificially created two sets of images of different qualities. They compared a highly homogenous set of faces (or butterflies) to the same set for which the faces and butterflies were changed in size and distorted (stretched and compressed along one axis) and randomly displaced off-center. Hence, the position of the most salient features was variable from trial to trial in the high variance condition, probably falling off the fovea in a large amount of trials and generating small and delayed N170s in individual trials, whereas it was constant and largely

---

[2] Note again that the methodology used by Thierry et al. being insufficiently described in their paper, it is impossible to know how much of the interstimulus correlations between their images is attributable to these irrelevant factors.

predictable in the low variance condition (see Fig. 6a of Thierry et al., 2007a).[3]

Considering these elements altogether, we suggest that the "physical variance" effect observed by Thierry et al. (reduced N170 in the high interstimulus variance condition) largely reflects a stimulus effect: a reduced visibility of category diagnostic features in the images from the "high variance" sets, because these features were displayed in an unusual view, presented out of the fovea[4], and/or distorted, causing a reduced N170 amplitude on these individual images.

To conclude this section, it is a basic methodological rule that in any paradigm such as the one used by Thierry et al., the exact same set of images should have been compared in the two conditions. For instance, one could have measured the N170 in response to 2 pictures of cars, being very different in size and viewpoint and show that the averaged N170 to each picture (e.g. 60 trials for picture 1, 60 trials for picture 2) is roughly identical. However, if a potential jitter factor was a real issue, averaging 30 trials of picture 1 with 30 trials of picture 2 should have led to a reduced N170. Another way to properly test for a potential role of interstimulus physical variance in partly accounting for the N170 amplitude would be to manipulate physical variance of a stimulus set parametrically, for instance by using morphed stimuli, and a well-balanced design. This could be combined with psychophysical measures of perceptual discrimination between stimuli and correlation measures with N170 amplitude. However, as we discuss below, such an experiment might not be very fruitful, given that the N170 amplitude appears to be largely driven by high-level visual processes that are independent of an interstimulus physical variance factor, even if it would be properly manipulated.

The respective role of low-level visual properties and high-level visual processes in accounting for N170 amplitude

In the present and following sections, we would like to go a little bit beyond Thierry et al.'s (2007a) study and discuss some general issues that are related to these authors' claim. Most importantly, because it has been raised in previous sections and it is a central issue in N170 face research, we would like to discuss the respective role of low-level visual properties of the stimulus versus high-level visual processes in accounting for the largest N170 amplitude in response to faces.

---

[3] What is even more problematic is the fact that the stimuli used in the high variance condition by Thierry et al. in their second experiment were distorted images. It may well be that these stimuli, which do not even respect the basic shape of a face, give rise to lower N170s than normal pictures of faces. Since butterflies probably have a shape that naturally varies much more, this stretching may not have affected the N170 to the same extent as for faces.

[4] Presumably, this effect of stimulus eccentricity did not affect the overall amplitude of preceding visual ERPs such as the P1 because what matters at that earlier stage of processing is not so much the diagnosticity of the features that are falling on the fovea, but their low-level properties. Moreover, whereas the N170 will be maximal for stimuli falling on the fovea (Rousselet et al., 2005; see also Jeffreys et al., 1992), the P1 can be decreased but also increased in amplitude for stimuli falling off the fovea, if they are presented in the upper or lower visual field for instance (Regan, 1989). Hence, averaging EEG segments to stimuli presented in the upper and lower visual fields (high variance) vs. centered stimuli (low variance) may not lead to observable amplitude difference on an averaged P1 component.

In most ERP studies of high-level visual processes, low-level properties of the categories of stimuli compared (e.g. energy in different frequency bands, local or global contrast), known to influence the amplitude of early visual potentials (see Regan, 1989), are usually not tightly controlled. In general, researchers should be aware that these factors may influence the amplitude and latency of visual ERPs, including the N170, and thus potentially affect the differential amplitude of this component for faces and nonface object categories. For instance, in the experiments of Thierry et al., in which object categories were compared to faces, low-level parameters between categories compared were not controlled, i.e., most notably the much higher contrast between upper and lower part of the stimuli for faces than cars or butterflies.

In some studies however, low-level properties (e.g. size, luminance, spatial frequency spectrum) have been controlled as much as possible between faces and the control object category compared (e.g. houses in Rousselet et al., 2005, 2007). In these conditions, the N170 is still much larger in response to faces, strongly indicating that low-level properties cannot account for the N170 face effect.

Although a careful control for low-level stimulus parameters is important to establish the validity of the early differences between object categories, in particular the larger response to face stimuli, we believe that such control is not always desirable in each and every ERP experiment. Indeed, when controlling for all low-level properties, the stimuli may become highly simplified or degraded, and the experiment may lose part of its ecological validity. Once it is established that an effect such as the larger electrophysiological N170 responses to faces than objects is not simply due to low-level visual properties left uncontrolled in a single study, it may be very interesting to test the response properties of the component with stimuli that present as many of the characteristics of natural faces. For instance, face stimuli may have relatively more power or may convey diagnostic information in different spatial frequency bands than object stimuli or visual scenes in general (Bosworth et al., 2006). While the spatial frequency power spectrum may not be a fundamental component of the definition of what constitutes a face stimulus for our perceptual system (the phase being critical), this factor may nevertheless play a role in the categorization of the stimulus as a face. For instance, it has been shown that photographs of human faces can be searched for efficiently (i.e., "pop out") among photographs of other objects (Hershler and Hochstein, 2005), but that this effect is mostly based on low-level factors such as the spatial frequency amplitude spectrum of face images (VanRullen, 2006). Obviously, this observation is interesting and should be taken into account when interpreting early differences between faces and objects. However, this certainly does not mean that in future experiments, one should control systematically for this factor and use only face and object stimuli with variations of image energy made completely irrelevant. Moreover, behavioral studies have shown that face perception is more sensitive to manipulations of spatial frequency content than object perception (Biederman and Kalocsai, 1997; Collin et al., 2004), perhaps because face perception relies on a broader range of information than object perception, encompassing edges, textures, and shading (Bruce and Humphreys, 1994). Equalizing pictures of faces and objects across spatial frequency bands may thus remove important information to process faces adequately.

Another way to ensure that early differences between object categories are not merely due to uncontrolled low-level parameters is to contrast the amplitude/latency difference between these stimuli in interaction with an identical stimulus manipulation. For instance,

if orientation affects the N170 for faces but not for objects, this effect is unlikely to be due to low-level differences between the categories compared (Rossion et al., 2000). More generally, the demonstration that category-related differences are not due to low-level visual properties comes from the accumulation of evidence within and between studies. That is, if the N170 is *systematically* larger in amplitude when compared to *many* different categories of objects that clearly differ according to low-level properties, it is a strong indication that the N170 effect cannot be accounted for by low-level visual properties of the stimuli.

In any event, there are plenty of observations in the literature indicating that the N170 amplitude is driven more by high-level processes (i.e., the perception of a face) than low-level visual features of the stimuli. One striking example is the finding that a stimulus perceived as a face when it is in the upright orientation, such as a two-tone "Mooney" (Mooney, 1956) face stimulus, elicits a decreased VPP/N170 amplitude when the exact same stimulus is presented upside-down and is no longer perceived as a face (Jeffreys, 1996; George et al., 2005; Latinus and Taylor, 2005; Fig. 9A). Similarly, the famous "Vegetable garden" painting by the Italian painter Giuseppe Arcimboldo, in which a face stimulus is
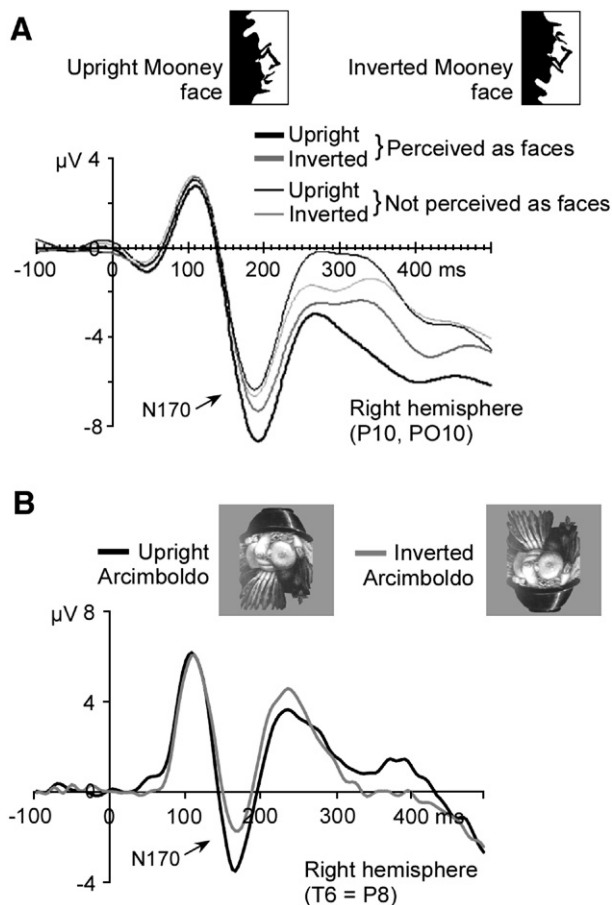

Fig. 9. (A) When two-tone ("Mooney") images are presented upright, they usually lead to the perception of a face stimulus, yielding a larger N170 than when the exact same pictures are presented inverted and do not lead to the perception of a face (figure adapted from George et al., 2005). (B) The same effect is observed for pictures of the paintings of G. Arcimboldo, where the face stimulus is perceived as emerging from the organization of nonface features such as fruits and vegetables (Caharel et al., in preparation).
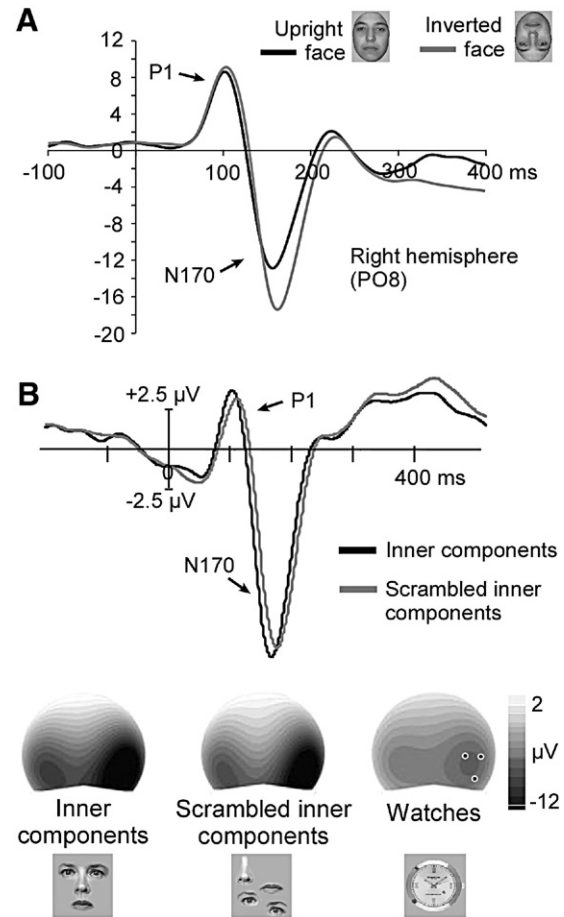

Fig. 10. (A) As demonstrated in numerous studies, inversion of a well segmented face stimulus, which leads to a massive decrement in individual recognition performance, causes a paradoxical substantial increase of N170 amplitude (e.g. Eimer, 2000b; Rossion et al., 1999; figure adapted from Jacques and Rossion, 2007a). Again, the stimuli compared are strictly identical in terms of global low-level properties (besides the phase of the stimulus). (B) Presenting faces with jumbled inner features, for which the interstimulus variance is extremely high compared to normal faces (Bentin et al., 2007b), does not reduce the N170 amplitude at all (data from Zion-Golumbic and Bentin, 2007) (see also George et al., 1996). The ERP waveforms are an average of 3 occipito-temporal electrodes in the right hemisphere (P8, PO8, P10) represented as small dots on the scalp topographies of the N170 (back-view of the head). The N170 topographies clearly show that whereas the N170 amplitude is similar for normally configured and jumbled inner features, the N170 is much smaller in response to pictures of watches.

perceived due to the combination of fruits and vegetables for instance, elicits a large N170. However, when the exact same stimulus is presented upside-down, the face is usually not perceived, leading to a reduced N170 (Caharel et al., in preparation; see Fig. 9B). These effects occur despite the fact that picture-plane inversion of a stimulus keeps constant all global low-level factors except for the phase of the stimulus.

When segmented faces are presented, they can be readily categorized as faces, even when presented upside-down. In these conditions, as illustrated in Fig. 10A, even though inversion of the face stimulus dramatically affects individual face recognition (Yin, 1969), it actually *increases* the N170 amplitude substantially (e.g. Bentin et al., 1996; Itier and Taylor, 2002, 2004a; Itier et al., 2006;

Jacques and Rossion, 2007a; Rossion et al., 1999, 2000, 2003; Rousselet et al., 2004a; Sagiv and Bentin, 2001). The nature of this paradoxical amplitude increase for inverted faces (as long as no visual noise is added to the face, see Schneider et al., 2007) is still unclear (see discussions in e.g. Bentin et al., 1996; Rossion et al., 1999; Itier et al., 2007) but it shows that keeping physical variance as well as low-level visual parameters between conditions equal, does not prevent the observations of large variations in N170 amplitude. In the same vein, the modulations of N170 amplitude to nonface objects induced by visual expertise (Tanaka and Curran, 2001), the reduction of the N170 amplitude to faces when objects of expertise are presented concurrently (Rossion et al., 2007, 2004), or the absence of N170 face effect in some congenital prosopagnosic individuals in contrast to normal controls (e.g. Bentin et al., 2007a) cannot be accounted for by low-level properties or by interstimulus physical variance factors, again because the exact same stimuli are compared. The same reasoning holds for the emergence of a N170 face-like response within session after conceptually priming the participants' awareness to the physiognomic value of stimuli, which are actually the same before and after priming (Bentin and Golland, 2002; Bentin et al., 2002). Finally, breaking the normal organization of facial features ("jumbling") to the point where interstimulus physical variance between stimuli is maximal does not reduce at all the N170 amplitude (Zion-Golumbic and Bentin, 2007; Fig. 10B) or may even increase it (George et al., 1996). All of these examples indicate that low-level properties and interstimulus physical variance may not account for much of the N170 amplitude variation.

In fairness, Thierry et al. did not argue that their observations could account for all of these interesting phenomena, which were not mentioned. However, in claiming that by manipulating a simple factor such as physical variance they could abolish a robust face-related effect that rather reflect high-level visual processes, they indirectly questioned the validity of all these observations of N170 modulations, which are highly relevant for understanding face categorization processes in the human brain.

P1 and N170: how early are preferential responses to faces observed?

While Thierry et al.'s abolished the N170 effect, they observed a much larger P1 – the central posterior component preceding the N170 (Figs. 5 and 7) – in response to faces than objects[5] and therefore claimed that (p. 509) "the P1 becomes de facto a better category sensitive ERP marker" (i.e., than the N170). Besides acknowledging that it may merely reflect differences in the low-level properties of the stimuli used (see previous section) the authors do not provide any satisfactory account of this observation.

Nevertheless, it brings the question of whether the P1 is truly larger in response to faces than objects in general, and if this effect is really larger and more consistent than the N170 face effect. More generally, how early in the visual processing stream are faces discriminated from other object categories?

The visual P1 (or P100) is an early component, peaking at around 100 ms following stimulus onset, and thought to originate from striate and extrastriate visual areas (e.g. Clark et al., 1995; Di Russo et al., 2002). It is known to be sensitive to many low-level properties of visual stimuli and is perhaps the most documented visual evoked potential.

Previous studies have reported larger P1 (or M1 in MEG) in response to faces than to objects (e.g. Eimer, 1998, 2000a; Goffaux et al., 2003; Itier and Taylor, 2004a; Liu et al., 2002), so the observation of Thierry et al. is neither very novel nor informative. The P1 may have also been larger to pictures of faces in some other studies, but this effect is not always analyzed (e.g. Eimer, 2000b; Itier et al., 2006; Rossion et al., 2000). However, such increased P1 amplitude for faces over objects is not consistently observed (e.g. Boutsen et al., 2006; Rossion et al., 2003; Rousselet et al., 2005, 2007) and is usually restricted to electrodes near the medial occipital pole[6] (e.g. Eimer, 1998; 2000a — see Figs. 5 and 7). Moreover, most studies, even those reporting P1 effects, find a much more robust difference between faces and objects at the level of the N170 (e.g. Goffaux et al., 2003; Itier and Taylor, 2004a; Liu et al., 2002; see Figs. 5 and 7). Importantly, the N170 effect is also observed in peak-to-peak measurements, i.e., if the N170 amplitude differences are measured with respect to P1 amplitude differences (e.g. Goffaux et al., 2003; Rossion et al., 2003). Unfortunately, Thierry et al. did not report peak-to-peak measurements of the N170 (i.e., subtracting P1 amplitudes from N170 values). They observed a significantly larger P1 for faces than objects but the following N170 was of equal amplitude for both categories. Therefore, their data actually indicate that the increase of amplitude taking place during the N170 time window may have been larger for faces vs. objects in their study as well (i.e., correcting for such P1 differences may have led to significant N170 effects; see Figs. 4b and 6b of Thierry et al., 2007a). In our short rebuttal of their study (Bentin et al., 2007b), we raised this point, expecting that Thierry and colleagues would provide these simple peak-to-peak analyses to qualify their claim, but the authors did not address it. This is again disappointing, in particular when considering that Thierry et al. measured the N170 at a majority of channels where the P1 is the largest on topographical maps (POZ, OZ, O1, O2, PO3, PO4) but where the N170 is small or non-existent (Figs. 5–7). Hence, a larger positive amplitude for faces over objects on the P1, as they found, complicates the interpretation of an (absence of) effect on the next component of opposite polarity, the N170. This is particularly true for electrodes PO3/4 and POZ at which the P1 is usually longer and delayed compared to occipital channels.

---

[5] The same caution than for the segmentation results in the N170 range must be taken regarding the segmentation results in the P1 range for several reasons. First, Thierry et al.'s finding of distinct maps for faces and objects in the P1 range is inconsistent with most – if not all – of the few previous studies that have used segmentation on face vs. nonface ERP analyses (Caldara et al., 2003; Itier and Taylor, 2004a; Rousselet et al., 2004b; Thierry et al., 2006). These studies did not find that different categories were associated with distinct segment maps in the P1 range, even though some of these studies (e.g. Itier and Taylor, 2004a) found amplitude difference between categories at the P1 similarly to Thierry et al. (2007a)'s finding. As for the N170, the interstimulus variance factor cannot be invoked to explain these discrepancies since the same maps were found in the high and low interstimulus variance conditions in Thierry et al. (2007a).

[6] This is entirely supported by the fact that while there is a P1 difference between faces and cars at occipital electrodes (O1/2) in the first experiment of Thierry et al. (2007a), the P1 category effect is no longer observed on low occipito-temporal electrodes (P7/8, PO9/10, see Fig. 1 of Thierry et al., 2007b), as is often observed in other studies (e.g. Eimer, 1998, 2000a; Fig. 5). Since early ERPs recorded at medial occipital sites arise mostly from lower level visual areas (Clark et al., 1995; Di Russo et al., 2002), this observation further supports the view that the P1 effect reported by Thierry et al. (2007a) and previous studies are likely to reflect differences in the low-level features between stimulus categories.

Note that we are not claiming here that peak-to-peak amplitude measurements should be carried out on a target component rather than baseline-to-peak measurements (see Regan, 1989; Picton et al., 2000 for discussions of this issue). We are not arguing either that Thierry et al. (2007a) should have concluded in favor of an N170 effect on the sole basis of a significant difference with a peak-to-peak analysis (i.e., without such an effect in baseline-to-peak measurements). However, when one observes a significant difference on the first of two such temporally close components (P1 and N170) that are maximal in the same EEG bandwidths (see Rousselet et al., 2007), the conclusions about amplitude modulations on the second component (or absence of) should be drawn with great care and tested by *both* baseline-to-peak and peak-to-peak measures. Ideally, the two measures should give congruent results, for instance a larger N170 to faces than objects whether an early P1 difference is found or not (e.g. Goffaux et al., 2003). Hence, by restricting their analysis to peak-to-baseline differences, Thierry et al. used a limited methodology to derive hasty and wrong conclusions.

Independently of the N170 sensitivity to faces, an issue that we clarified in the present manuscript, do P1 effects also reflect early face categorization processes? Because previous studies have indeed reported larger P1 in response to faces than to objects (e.g. Eimer, 1998, 2000a; Goffaux et al., 2003; Itier and Taylor, 2004a; Liu et al., 2002), it is fair to admit that there are probably electrophysiological differences between faces and nonface object categories taking place earlier than the N170 onset (130 ms), that is, at the level of the P1. However, one must remain extremely cautious in overinterpreting these P1 effects, at least for two reasons. First, as indicated earlier, the P1 is an early visual ERP highly sensitive to low-level visual parameters such as luminance, color, contrast or spatial frequencies of the stimulus (see Regan, 1989). Small P1 (or M1) face effects, inconsistently found in the literature, may be due to differences in the global or local low-level parameters between stimulus categories generally left uncontrolled in ERP and MEG studies. It may also be observed in some experiments in which the signal-to-noise ratio is extremely poor, which may lead to an absence of visible components in object conditions (see e.g. Fig. 1 of Liu et al., 2002). For instance, an early MEG study showed that the early M1 difference between photographs of faces and other categories was cancelled when removing surface information (color and texture) of the face stimuli (Halgren et al., 2000), prompting the authors to attribute this early difference to low-level properties (the M170 effect remained conspicuous). Other studies in which the faces and object stimuli were carefully controlled do not report P1 differences between categories (e.g. Rousselet et al., 2005). A second reason for remaining cautious about early P1 effects is that, as noted above, unlike N170 effects, these categorical differences at the level of the P1 are simply not consistently found. For instance, we found both a larger P1 and N170 amplitude to pictures of faces than cars for 3/4 views of the two categories (Fig. 7). However, when we compared symmetrical images, that is, *full-front* pictures of faces and cars, there was no P1 effect at all: the larger amplitude to faces was restricted to the N170 time window (Rossion and Jacques, in preparation; Fig. 7). This observation further shows, if necessary, that the claim made by Thierry et al. that "the P1 becomes de facto a better category sensitive ERP marker" (than the N170) is unfounded. The P1 may perhaps be an *earlier*, although unreliable, marker of the face-preferential response than the N170, but certainly *not* a *better* marker of face-specific processes.

In addition to showing a stronger and more consistent face-preferential response, the N170 effect is also well known for its right lateralized topographical distribution on occipito-temporal sites (Bentin et al., 1996; Rossion et al., 2003), which is rarely the case for the P1 differences (Fig. 5). The N170 effect is thus consistent with the dominant role of the right occipito-temporal cortex in face perception, as derived from lesion analyses of prosopagnosia (e.g. Sergent and Signoret, 1992; Bouvier and Engel, 2006; Sorger et al., 2007) and neuroimaging studies in the normal brain (e.g. Sergent et al., 1992; Kanwisher et al., 1997).

If we leave aside for a forthcoming short paragraph the question of face categorical sensitivity, there are many other reasons why the N170 is favored over the P1 by ERP researchers interested in face processing. For instance, whereas raw effects of face inversion (amplitude and latency increases) can be observed on the P1 (e.g. Itier and Taylor, 2002, 2004a), these effects are much stronger and more consistent on the N170 (e.g. Eimer, 2000b; Itier and Taylor, 2002, 2004a; Linkenkaer-Hansen et al., 1998; Rebai et al., 2001; Rossion et al., 1999, 2000, 2003; Rousselet et al., 2004a; Sagiv and Bentin, 2001) (see Fig. 10). Most importantly, electrophysiological effects are not correlated to the behavioral effects of face inversion before the N170 time window (see Jacques and Rossion, 2007a). A functional dissociation between the P1 (or M1) and the N170/M170 has also been found by studies varying the amount of visual noise or the noise spatial frequency added to a face image. For instance, the P1/M1 is strongly correlated with the amount of noise in an image (Tarkiainen et al., 2002) or the noise's spatial frequency (Tanskanen et al., 2005) but is not correlated with the amount of face information in the image (Jemel et al., 2003; Tanskanen et al., 2005; Tarkiainen et al., 2002). By contrast, in the same studies, the amplitude and latency of the N170/M170 are strongly correlated with the perception of a face stimulus. In short, many observations indicate that fundamental face processes appear to take place primarily during the N170 time window rather than the P1.
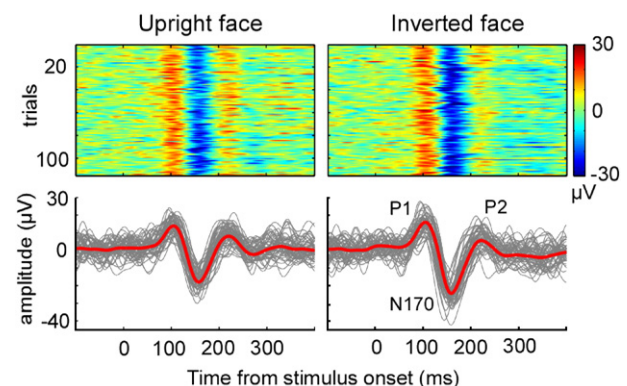

Fig. 11. This figure shows that the N170 response is best described as a strong increase of power time-locked and phase-locked to stimulus onset. *Upper row* depicts ERP images for 115 EEG single trials time-locked to the onset of a stimulus (no vertical smoothing was applied in the ERP image). Note the strong increase of amplitude during P1, N170 and P2 components relative to the 100 ms before stimulus onset. Bottom row shows an average of the trials depicted in each ERP image. The figure also illustrates that difference between type of stimuli (here upright and inverted faces; left and right column respectively) at the N170 is also accounted for by a change in amplitude in each individual trial. This is also true for the N170 face effect (Rousselet et al., 2007). *Bottom row* shows individual trials (thin gray traces) and an average of all trials (thick red trace). This also shows the increase of amplitude during the time windows of the P1, N170 and P2 components.

In this section, we would also like to briefly address the important question raised at the beginning of the section: how early in the visual processing stream are faces discriminated from other object categories? Given the finding of a larger P1 (~80 ms onset, maximal at ~110 ms) for faces in some studies, is the N170 time window (~130 ms onset) too late? The question is not an easy one to answer, and in fact it is a question that may not have a single and simple answer. For comparison, neurons recorded in the monkey infero-temporal cortex start firing selectively to faces at around 80–90 ms (e.g. Baylis et al., 1987; Matsumoto et al., 2005), with the earliest and mean onset latencies being observed at around 70 and 100 ms respectively (e.g. Baylis et al., 1987; Kiani et al., 2005). Considering that these neurons will fire slightly later in the bigger human brain (see Foxe and Schroeder, 2005; Schroeder et al., 1998, 2004), selective response to faces starting at around 120–130 ms, at the level of the N170 onset, seems to be a reasonable time frame (with a maximal discriminability between faces and object categories at the 160–170 ms peak).

However, as already mentioned, faces differ from other object categories on multiple low-level visual properties and are likely to elicit differential responses earlier in the hierarchy of cortical areas of the visual stream. Hence, in most situations, there is probably sufficient evidence accumulated in early visual areas to perform a face/nonface categorization task above chance level *before* the N170 onset latency. For instance, ERP and forced-choice saccadic eye movement studies indicate that categorization of animal pictures in visual scenes can take place within 120–150 ms following stimulus onset, including the perceptual decision (Kirchner and Thorpe, 2006; Thorpe et al., 1996; VanRullen and Thorpe, 2001). These latencies suggest that visual stimuli may be, under certain circumstances, categorized reliably as faces before the N170 onset, perhaps within the time window of the P1. However, while face categorization reflected on the N170 can be largely independent of low-level cues, rather reflecting what the observer is seeing (e.g. Bentin and Golland, 2002; Bentin et al., 2002; George et al., 2005; Jeffreys, 1996; Sagiv and Bentin, 2001, see Fig. 9), earlier effects may rather arise because of the presence of such low-level cues that are statistically more frequently associated with faces (e.g. contrast, spatial frequency content, color distribution). Supporting this low-level visual account, when the location and size of faces in visual scenes are largely unpredictable, the fastest behavioral responses to detect a face are largely unaffected by stimulus vertical inversion (Rousselet et al., 2003).

More generally, the speed at which faces are categorized by the visual system, whether it is slightly before or at the N170 onset, is likely to depend on the nature of other stimuli, either presented simultaneously (in a visual scene) or interleaved in a face categorization task. That is, there is no single answer to the question of the speed at which faces can be categorized by the visual system: it depends on the rate of accumulation of evidence in multiple distributed populations of neurons in the visual system and is dependent on the categorization context (distracters, visual context, subjects expectations…). Under certain circumstances, visual stimuli may indeed be categorized as faces before the N170 onset, but this is likely to be due to low-level properties that are statistically associated with face photographs rather than the perception of a face per se. Moreover, given the early onset latency of the P1 (~80 ms) as compared to the mean onset latency of face-selective neurons (100 ms in the monkey brain, Kiani et al., 2005, probably slightly later in the human brain), it is unlikely that this type of categorization based on image statistics involves 'face-specific'

neural processes (i.e., the processes carried out by these face neurons).

To conclude this section, we would like to reiterate that the occipito-temporal N170 offers a more reliable time frame (130–200 ms) than earlier potentials such as the P1 to investigate the nature of face perceptual mechanisms in the human brain. In any event, future studies should go beyond a mere debate between ERP components (P1 vs. N170), which should not be interpreted as reflecting fixed stages of face processing. For instance, by performing point-by-point correlation between electrophysiological and behavioral responses, the exact time point at which the effect of stimulus and task manipulations arise can be identified (e.g. Jacques and Rossion, 2007a; Philiastides and Sajda, 2006), analyses which may reveal significant effects *after* the P1 component, in the downward slope of the N170 (e.g. see Jacques and Rossion, 2007a).

Why is the N170 larger in amplitude to faces?

The N170 is larger in amplitude when faces are presented relative to other object categories. This effect cannot be accounted for by a simple factor such as a lower physical variance of face stimuli as compared to other objects. Furthermore, it is unlikely that a greater *perceptual* similarity of face stimuli would even account for substantial part of this phenomenon. Why then is the N170 larger for faces and why should we care after all? How would this help us understand how faces are processed in the human brain?

Interpreting a differential amplitude of a scalp ERP component between two conditions is not straightforward because it can be related to many underlying neural events (see Luck, 2005). To simplify, there are mostly two ways to account for the larger N170 to faces than objects as revealed by averaging EEG segments time-locked to stimulus onset. According to the first traditional view, ERP components originate from a massive synchronized increase of post-synaptic neural activity time-locked and phase-locked to stimulation onset, superimposed onto background electrophysiological activity unrelated to the stimulation (the signal+noise model). In this framework, the larger N170 to faces could simply reflect a stronger increase of neural activity as compared to objects, in visual areas from which this component originates. This would be reflected as a stronger increase of power at a fixed latency and polarity at the scalp level for face stimuli (Fig. 11).

According to a second view, the larger N170 to faces may also be due to face stimuli eliciting an ERP response at a more consistent latency from trial to trial as compared to objects, without an increase of amplitude. This latest account would be compatible with the speculation of Thierry et al. (2007a) that a smaller N170 for objects would simply result from a larger time jitter in the peak of the N170 from trial to trial in response to these stimuli. As for face stimuli, this decrease in inter-trial jitter would correspond either to a lower variance in the peak latency of the N170 from trial to trial or to a more precise phase resetting of ongoing EEG oscillations (i.e., preceding the stimulus). Phase resetting in EEG is a realignment of the EEG ongoing oscillations due to the stimulation, without any substantial increase of amplitude, leading to a visible (time-locked) ERP component after averaging in the time domain (Sayers and Beagley, 1974). The larger amplitude of the N170 to faces could then be due to a more precise temporal alignment of the EEG phase from trial to trial in response to faces compared to nonface objects.

Interestingly, it has been proposed that the visual N1 component to simple stimuli can indeed be largely generated by such a phase resetting of EEG ongoing oscillations in the alpha range (Makeig et

al., 2002). Thus, given the nature of their claim, Thierry et al. (2007a) could have searched for a better support to these hypotheses, by measuring the phase-locking factors[7] for their different conditions across time (e.g. TallonBaudry et al., 1996). However, there is no evidence in favor of the pure phase resetting model for the N170 face effect: the largest N170 to faces is associated with a massive increase of power in the 5- to 15-Hz band time-locked to stimulus onset (Henson et al., 2005; Rousselet et al., 2007). There is also an increase in phase resetting at that latency, but which appears to be due to the effect of the time-locked power increase on the EEG phase rather than to a pure phase resetting (Rousselet et al., 2007; see Sauseng et al., 2007 for a discussion of this issue).[8]

Even though the larger N170 to faces than other categories is due to a fixed increase of power in localized brain areas between 130 and 200 ms,[9] the interpretation of this amplitude difference remains ambiguous with respect to the underlying sources, as originally discussed by several authors (Bentin et al., 1996; Botzel et al., 1995; Jeffreys, 1996). That is, is the N170 due to the addition of multiple cortical sources, i.e., several sub-components, or to a major/dominant source that would be of stronger magnitude for faces than objects? Source localization studies of the face N170/M170 with various methods have given rather discrepant results (e.g. bilateral sources in the posterior fusiform gyrus/lateral occipito-temporal complex: Deffke et al., 2007; Henson et al., 2007; Rossion et al., 2003; Schweinberger et al., 2002; in the lateral parts of the temporal cortex: Shibata et al., 2002; Watanabe et al., 2003; in the superior temporal sulcus: Itier and Taylor, 2004b; in bilateral fusiform gyri, together with multiple activation in a parieto–temporal–occipital network of areas: Herrmann et al., 2005). These results are strongly indicative of the presence of multiple cortical sources interlocked in time (130–200 ms) accounting for the N170 component and its larger amplitude to faces, in line with fMRI recordings during the same face experiments (e.g. Henson et al., 2003; Horovitz et al., 2004). Neuroimaging studies (e.g. Sergent et al., 1992; Haxby et al., 2000) as well as intracranial recordings in the human brain (e.g. Allison et al., 1999) also indicate that some of these sources may be preferentially or even selectively activated to faces. Hence, the larger N170 component observed at the surface of the brain in response to faces appears to be due to the *additional* contribution of local sources at that latency that are face-selective.

According to the view of the N170 as originating from multiple sources, modulations of amplitude on the scalp during the N170 time window represent a mixture of underlying neural activity in high-level visual areas. The time window of the component is of particular interest for face processing precisely because it appears to capture a large part of the early *selective* visual responses to face stimuli that take place in the occipito-temporal cortex at a finer spatial scale. Thus, it allows researchers to test non-invasively whether perceptual face processes can be modulated on line by a number of factors (stimulus manipulations, task and attentional factors, see the Introduction section) and to investigate the nature of the early perceptual face representations. In particular, the larger N170 to faces is a starting point to address the question of whether these earlier processes reflect the activity of fixed underlying populations of neurons devoted to faces, or if these neurons can be potentially recruited for other visual categories, for instance following visual expertise training (e.g. Rossion et al., 2004, 2007). Thus, while the spatial resolution of scalp ERP (or even MEG components) is poor compared to fMRI, studies of the face-related N170 remain invaluable to investigate the nature and the dynamics of face processes in the human brain, for the method offers excellent time resolution while sampling data from the whole system simultaneously.

Summary and conclusions: ten lessons on the N170 from an unfortunate publication

Thierry and colleagues (2007a) recently reported a study that claims to eliminate the larger N170 amplitude to faces by controlling for a methodological artifact of interstimulus physical variance. A careful reading of this study, contrasted with the existing N170 literature, shows that:

1. These authors based their study on a crude and inadequate definition of interstimulus variance among pictures (inter-pixel correlation), which should not be referred to as "perceptual" but "physical", and is not necessarily relevant to future ERP studies of the N170 that compare stimuli from different categories.

2. Many published studies actually controlled for physical variance as defined by Thierry et al. and reported a larger N170 to pictures of faces than objects, including highly familiar categories such as cars. Hence, the claim that a smaller physical variance among face stimuli in previous studies accounts for the N170 effect is incorrect. It is also fundamentally incompatible with an absence of consistent N170 peak delay and smearing of the component for nonface objects, which should have been observed if there was an increase in trial-to trial jitter relative to faces.

3. In contrast to many previous studies, Thierry et al. did not control for physical interstimulus variance when comparing between faces and nonface objects in their own experiments. They only manipulated this factor, in a non-balanced design. Only a few previous studies may suffer from the same limitations, most notably experiments that rely on a so-called 'face localizer' contrast (comparing pictures of faces to objects from multiple categories) and a previous ERP study of the authors themselves (Thierry et al., 2006).

4. While the N170 is most prominent and usually analyzed on lower occipito-temporal channels, Thierry et al. measured the amplitude of this component on the wrong electrode sites. They did not include some of the most sensitive electrode sites for the N170, and they displayed comparisons between

---

[7] The phase-locking factor is a measure across time points and frequency of the precision of the EEG phase alignment from trial to trial. It ranges from 1 (EEG at a given time point perfectly in phase from trial to trial) to 0 (No relationship between the phase of the EEG from trial to trial at a given time point).

[8] As a matter of fact, observations of individual trials in the data reported by Makeig and collaborators (2002) also indicate a massive increase of power time-locked to stimulus onset, which may in fact be largely responsible in itself for the generation of the N1 to simple visual stimuli in that study also (see Yeung et al., 2004). Evidence for pure phase resetting as accounting for evoked potentials is a complicated issue because of the influence of evoked components on the EEG phase (see Sauseng et al., 2007).

[9] This does not exclude that a differential phase resetting may potentially account for other amplitude modulations of the face-related N170, such as the increased amplitude for inverted faces, which, despite being physically identical, are more perceptually similar to each other than upright faces. There is, however, no current evidence in support of this suggestion.

faces and objects at less sensitive channels. Worse, they included a set of electrodes located too high and medially on the scalp (e.g. OZ, POZ, PO3, PO4). These electrodes are never included in N170 measurements (especially when comparing faces to other objects) and were not considered in a previous study by the same authors. In doing so, they largely missed the N170 effect rather attributing too much weight to the preceding positive potential (P1) that was larger to faces than objects in their study and to the vertex positive counterpart (VPP) of the N170. Hence, they likely cancelled out any category effect on the N170 by merging channels showing a larger amplitude for faces, but of opposite polarities on the scalp. Presumably, this bias in channel selection was not apparent because no report was made of putative interactions between electrode and the independent variables.

5. Rather than manipulating interstimulus physical variance in a factorial design by comparing the same sets of images, Thierry et al. compared sets of images of different quality in their experiments. The N170 amplitude difference between their low and high physical variance conditions is likely to be due to a lower visibility of category diagnostic features for the images in the "high variance" set because they are masked (unusual viewpoints), presented out of fovea, and/or distorted which causes a reduced N170 amplitude for each individual image in this condition.

6. In general, the N170 face effect cannot be accounted for by low-level visual properties of the stimuli. Even though such a control is important to establish the validity of the N170 effect, degrading face and object stimuli in order to control for all low-level parameters is not always desirable, both for ecological reasons and because faces are usually associated with low-level properties that are an integral part of what constitutes a face stimulus for our visual system.

7. The claim that the larger N170 to faces can be accounted for by uncontrolled physical variance or low-level properties is inconsistent with multiple amplitude variations observed on this component following stimulus and task/context manipulations, e.g. picture-plane inversion, visual expertise, priming or attention or other manipulations in which physical variance and low-level parameters are kept constant between the conditions compared. All these effects indicate that, unlike the P1, the N170 amplitude is rather driven by high-level visual processes than low-level properties. Moreover, some of these effects are larger for faces as compared to nonface stimuli (e.g. inversion effect), indicating that the N170 is particularly sensitive to faces.

8. In some studies, there are significant differences between faces and other object categories preceding the N170 onset (~130 ms), i.e., at the level of the posterior P1 component, suggesting that the visual system may have accumulated enough evidence to categorize reliably a visual stimulus as a face before 130 ms. However, these early effects are inconsistently found and largely driven by differences in low-level properties, unlike the categorical differences found during the N170 time window. Furthermore, when amplitude differences are found on the preceding P1 visual component, it is recommended to measure the N170 both with respect to the baseline signal (before stimulus onset) and also to the preceding P1 peak. Consistent results between the two measures (peak-to-peak and peak-to-baseline) are recommendable for deriving conclusions about the N170 effects.

9. The occipito-temporal N170 shows reliable differences between faces and other object categories (which also differ among each other) between 130 and 200 ms following stimulus onset. While Thierry et al. (2007a) did not take hemispheric differences into account when analyzing their data, this difference is most significant in the right hemisphere, consistent with evidence from multiple sources, including neuroimaging and lesion studies of face processing. This time frame is compatible with average onset latency of face-selective cells in the monkey's infero-temporal cortex (about 100 ms), even though the N170 component itself in humans likely originates from multiple cortical sources interlocked in time between 130 and 200 ms. The larger amplitude of the N170 potential on the scalp to faces may simply reflect the additional contribution of a large subset of neural sources devoted selectively to faces.

10. The N170 corresponds to a time-locked and phase-locked increase of EEG power between 5 Hz and 15 Hz rather than to a phase resetting of an ongoing oscillation without increase of power. Similarly, there is no evidence that the larger N170 to faces than objects would be due to an increase in phase resetting between trials as compared to objects (larger inter-trial jitter), in complete contradiction with Thierry et al.'s speculations.

To conclude, Thierry et al. created a problem about the N170 face effect that did not exist in previous studies and was incompatible with reported evidence. This claim, reported in a wide scope journal, is misleading for a large audience of Cognitive Neuroscientists. In particular, it is important that the reader does not misunderstand Thierry et al.'s claim as being related to the debate about the N170 face effect reflecting modular processes for faces. Researchers in this field do not argue about the larger N170 response to faces than objects. This is considered as an established fact, just as everyone agrees that there are several high-level visual areas that respond more to faces than objects in the human brain. A matter of debate is indeed whether these preferential responses reflect domain-specific (modular) processes or rather visual processes that remain plastic in the adult brain and can be potentially recruited for nonface objects, for instance following visual expertise training. However, this debate has nothing to do with an uncontrolled artifact such as interstimulus physical variance. In contrast to Thierry et al.'s report, flawed by numerous methodological and conceptual faults, the literature reviewed above indicates that the special status that the N170 enjoys in the face literature is entirely justified: this component still has an extremely interesting future as a critical time window for investigating human face processing.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2007.10.011.

References

Allison, T., Puce, A., Spencer, D.D., McCarthy, G., 1999. Electrophysiological studies of human face perception. I: Potentials generated in occipitotemporal cortex by face and non-face stimuli. Cereb. Cortex 9, 415–430.

Baylis, G.C., Rolls, E.T., Leonard, C.M., 1987. Functional subdivisions of the temporal lobe neocortex. J. Neurosci. 7, 330–342.

Bentin, S., Deouell, L.Y., 2000. Structural encoding and identification in face processing: ERP evidence for separate mechanisms. Cogn. Neuropsychol. 17, 35–54.

Bentin, S., Golland, Y., 2002. Meaningful processing of meaningless stimuli: the influence of perceptual experience on early visual processing of faces. Cognition 86, B1–B14.

Bentin, S., McCarthy, G., Perez, E., Puce, A., Allison, T., 1996. Electrophysiological studies of face perception in humans. J. Cogn. Neurosci. 8, 551–565.

Bentin, S., Sagiv, N., Mecklinger, A., Friederici, A., von Cramon, Y.D., 2002. Priming visual face-processing mechanisms: electrophysiological evidence. Psychol. Sci. 13, 190–193.

Bentin, S., DeGutis, J.M., D'Esposito, M., Robertson, L.C., 2007a. Too many trees to see the forest: performance, event-related potential, and functional magnetic resonance imaging manifestations of integrative congenital prosopagnosia. J. Cogn. Neurosci. 19, 132–146.

Bentin, S., Taylor, M.J., Rousselet, G.A., Itier, R.J., Caldara, R., Schyns, P.G., Jacques, C., Rossion, B., 2007b. Much ado about nothing: controlling for interstimulus perceptual variance does not abolish N170 face sensitivity. Nat. Neurosci. 10, 801–802.

Biederman, I., Kalocsai, P., 1997. Neurocomputational bases of object and face recognition. Philos. Trans. R. Soc. Lond., B Biol. Sci. 352, 1203–1219.

Blanz, V., Tarr, M.J., Bulthoff, H.H., 1999. What object attributes determine canonical views? Perception 28, 575–599.

Bosworth, R.G., Bartlett, M.S., Dobkins, K.R., 2006. Image statistics of American sign language: comparison to faces and natural scenes. J. Opt. Soc. Am. 23, 2085–2096.

Botzel, K., Grusser, O.J., 1989. Electric brain potentials-evoked by pictures of faces and non-faces—a search for face-specific EEG-potentials. Exp. Brain Res. 77, 349–360.

Botzel, K., Schulze, S., Stodieck, S.R.G., 1995. Scalp topography and analysis of intracranial sources of face-evoked potentials. Exp. Brain Res. 104, 135–143.

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.

Bouvier, S.E., Engel, S.A., 2006. Behavioral deficits and cortical damage loci in cerebral achromatopsia. Cereb. Cortex 16, 183–191.

Bruce, V., Humphreys, G.W., 1994. Recognizing objects and faces. Vis. Cogn. 141–180.

Caharel, S., Lalonde, R., Bernard, C., Leleu, A., Rebaï, M., in preparation. The N170: an electrophysiological index of visual categorization dependent on the level of familiarity with objects and face formats.

Caldara, R., Thut, G., Servoir, P., Michel, C.M., Bovet, P., Renault, B., 2003. Face versus non-face object perception and the 'other-race' effect: a spatio-temporal event-related potential study. Clin. Neurophysiol. 114, 515–528.

Carmel, D., Bentin, S., 2002. Domain specificity versus expertise: factors influencing distinct processing of faces. Cognition 83, 1–29.

Clark, V.P., Fan, S., Hillyard, S.A., 1995. Identification of early visual evoked potential generators by retinotopic and topographic analyses. Hum. Brain Mapp. 2, 170–187.

Collin, C.A., Liu, C.H., Troje, N.F., McMullen, P.A., Chaudhuri, A., 2004. Face recognition is affected by similarity in spatial frequency range to a greater degree than within-category object recognition. J. Exp. Psychol. Hum. Percept. Perform. 30, 975–987.

Damasio, A.R., Damasio, H., Van Hoesen, G.W., 1982. Prosopagnosia: anatomic basis and behavioral mechanisms. Neurology 32, 331–341.

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

Di Russo, F., Martinez, A., Sereno, M.I., Pitzalis, S., Hillyard, S.A., 2002. Cortical sources of the early components of the visual evoked potential. Hum. Brain Mapp. 15, 95–111.

Eimer, M., 1998. Does the face-specific N170 component reflect the activity of a specialized eye processor? NeuroReport 9, 2945–2948.

Eimer, M., 2000a. Attentional modulations of event-related brain potentials sensitive to faces. Cogn. Neuropsychol. 17, 103–116.

Eimer, M., 2000b. Effects of face inversion on the structural encoding and recognition of faces—Evidence from event-related brain potentials. Cogn. Brain Res. 10, 145–158.

Faust, C., 1955. Evaluation of posttraumatic organic personality disorders. Dtsch. Med. Wochenschr. 80, 1237–1239.

Foxe, J.J., Schroeder, C.E., 2005. The case for feedforward multisensory convergence during early cortical processing. NeuroReport 16, 419–423.

Galton, F., 1883. Inquiries into human faculty and its development. Macmillan, London.

Gauthier, I., Behrmann, M., Tarr, M.J., 1999. Can face recognition really be dissociated from object recognition? J. Cogn. Neurosci. 11, 349–370.

George, N., Evans, J., Fiori, N., Davidoff, J., Renault, B., 1996. Brain events related to normal and moderately scrambled faces. Brain Res. Cogn. Brain Res. 4, 65–76.

George, N., Jemel, B., Fiori, N., Chaby, L., Renault, B., 2005. Electrophysiological correlates of facial decision: Insights from upright and upside-down Mooney-face perception. Cogn. Brain Res. 24, 663–673.

Goffaux, V., Gauthier, I., Rossion, B., 2003. Spatial scale contribution to early visual differences between face and object processing. Cogn. Brain Res. 16, 416–424.

Halgren, E., Raij, T., Marinkovic, K., Jousmaki, V., Hari, R., 2000. Cognitive response profile of the human fusiform face area as determined by MEG. Cereb. Cortex 10, 69–81.

Haxby, J.V., Hoffman, E.A., Gobbini, M.I., 2000. The distributed human neural system for face perception. Trends Cogn. Sci. 4, 223–233.

Henson, R.N., Goshen-Gottstein, Y., Ganel, T., Otten, L.J., Quayle, A., Rugg, M.D., 2003. Electrophysiological and haemodynamic correlates of face perception, recognition and priming. Cereb. Cortex 13, 793–805.

Henson, R.N., Kiebel, S., Kilner, J., Friston, K., Hillebrand, A., Barnes, G.R., Singh, K.D., 2005. Time–frequency SPMs for MEG data on face perception: power changes and phase-locking. HBM Abstract.

Henson, R.N., Mattout, J., Singh, K.D., Barnes, G.R., Hillebrand, A., Friston, K., 2007. Population-level inferences for distributed MEG source localisation under multiple constraints: application to face-evoked fields. NeuroImage 38, 422–438.

Herrmann, M.J., Ehlis, A.C., Muehlberger, A., Fallgatter, A.J., 2005. Source localization of early stages of face processing. Brain Topogr. 18, 77–85.

Hershler, O., Hochstein, S., 2005. At first sight: a high-level pop out effect for faces. Vis. Res. 45, 1707–1724.

Holmes, A., Vuilleumier, P., Eimer, M., 2003. The processing of emotional facial expression is gated by spatial attention: evidence from event-related brain potentials. Brain Res. Cogn. Brain Res. 16, 174–184.

Horovitz, S.G., Rossion, B., Skudlarski, P., Gore, J.C., 2004. Parametric design and correlational analyses help integrating fMRI and electrophysiological data during face processing. NeuroImage 22, 1587–1595.

Itier, R.J., Taylor, M.J., 2002. Inversion and contrast polarity reversal affect both encoding and recognition processes of unfamiliar faces: a repetition study using ERPs. NeuroImage 15, 353–372.

Itier, R.J., Taylor, M.J., 2004a. N170 or N1? Spatiotemporal differences between object and face processing using ERPs. Cereb. Cortex 14, 132–142.

Itier, R.J., Taylor, M.J., 2004b. Source analysis of the N170 to faces and objects. NeuroReport 15, 1261–1265.

Itier, R.J., Latinus, M., Taylor, M.J., 2006. Face, eye and object early processing: what is the face specificity? NeuroImage 29, 667–676.

Itier, R.J., Alain, C., Sedore, K., McIntosh, A.R., 2007. Early face processing specificity: it's in the eyes! J. Cogn. Neurosci. 19, 1815–1826.

Jacques, C., Rossion, B., 2004. Concurrent processing reveals competition between visual representations of faces. NeuroReport 15, 2417–2421.

Jacques, C., Rossion, B., 2007a. Early electrophysiological responses to multiple face orientations correlate with individual discrimination performance in humans. NeuroImage 36, 863–876.

Jacques, C., Rossion, B., 2007b. Electrophysiological evidence for temporal dissociation between spatial attention and sensory competition during human face processing. Cereb. Cortex 17, 1055–1065.

Jasper, H.H., 1958. The ten–twenty system of the International Federation. Electroencephalogr. Clin. Neurophysiol. 10, 371–375.

Jeffreys, D.A., 1989. A face-responsive potential recorded from the human scalp. Exp. Brain Res. 78, 193–202.

Jeffreys, D.A., 1996. Evoked potential studies of face and object processing. Vis. Cogn. 3, 1–38.

Jeffreys, D.A., Tukmachi, E.S.A., Rockley, G., 1992. Evoked-potential evidence for human brain mechanisms that respond to single fixated faces. Exp. Brain Res. 91, 351–362.

Jemel, B., Schuller, A.M., Cheref-Khan, Y., Goffaux, V., Crommelinck, M., Bruyer, R., 2003. Stepwise emergence of the face-sensitive N170 event-related potential component. NeuroReport 14, 2035–2039.

Joyce, C., Rossion, B., 2005. The face-sensitive N170 and VPP components manifest the same brain processes: the effect of reference electrode site. Clin. Neurophysiol. 116, 2613–2631.

Joyce, C.A., Schyns, P.G., Gosselin, F., Cottrell, G.W., Rossion, B., 2006. Early selection of diagnostic facial information in the human visual cortex. Vis. Res. 46, 800–813.

Kanwisher, N., McDermott, J., Chun, M.M., 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. J. Neurosci. 17, 4302–4311.

Kiani, R., Esteky, H., Tanaka, K., 2005. Differences in onset latency of macaque inferotemporal neural responses to primate and non-primate faces. J. Neurophysiol. 94, 1587–1596.

Kirchner, H., Thorpe, S.J., 2006. Ultra-rapid object detection with saccadic eye movements: visual processing speed revisited. Vis. Res. 46, 1762–1776.

Kovacs, G., Zimmer, M., Banko, E., Harza, I., Antal, A., Vidnyanszky, Z., 2006. Electrophysiological correlates of visual adaptation to faces and body parts in humans. Cereb. Cortex 16, 742–753.

Latinus, M., Taylor, M.J., 2005. Holistic processing of faces: learning effects with Mooney faces. J. Cogn. Neurosci. 17, 1316–1327.

Linkenkaer-Hansen, K., Palva, J.M., Sams, M., Hietanen, J.K., Aronen, H.J., Ilmoniemi, R.J., 1998. Face-selective processing in human extra-striate cortex around 120 ms after stimulus onset revealed by magneto- and electroencephalography. Neurosci. Lett. 253, 147–150.

Liu, C.H., Chaudhuri, A., 2002. Reassessing the 3/4 view effect in face recognition. Cognition 83, 31–48.

Liu, J., Harris, A., Kanwisher, N., 2002. Stages of processing in face perception: an MEG study. Nat. Neurosci. 5, 910–916.

Liu, J., Higuchi, M., Marantz, A., Kanwisher, N., 2000. The selectivity of the occipitotemporal M170 for faces. NeuroReport 11, 337–341.

Luck, S., 2005. An Introduction to the Event-Related Potential Technique. MIT Press, Cambridge, MA.

Makeig, S., Westerfield, M., Jung, T.P., Enghoff, S., Townsend, J., Courchesne, E., Sejnowski, T.J., 2002. Dynamic brain sources of visual evoked responses. Science 295, 690–694.

Matsumoto, N., Okada, M., Sugase-Miyamoto, Y., Yamane, S., Kawano, K., 2005. Population dynamics of face-responsive neurons in the inferior temporal cortex. Cereb. Cortex 15, 1103–1112.

Mooney, C.M., 1956. Closure with negative after-images under flickering light. Can. J. Psychol. 10, 191–199.

Nederhouser, M., Yue, X., Biederman, I., 2006. Predicting psychophysical similarity of complex shapes from measures of physical similarity. J. Vis. 6, 320.

Philiastides, M.G., Sajda, P., 2006. Temporal characterization of the neural correlates of perceptual decision making in the human brain. Cereb. Cortex 16, 509–518.

Picton, T.W., Bentin, S., Berg, P., Donchin, E., Hillyard, S.A., Johnson, R., Miller, G.A., Ritter, W., Ruchkin, D.S., Rugg, M.D., Taylor, M.J., 2000. Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria. Psychophysiology 37, 127–152.

Rebai, M., Poiroux, S., Bernard, C., Lalonde, R., 2001. Event-related potentials for category-specific information during passive viewing of faces and objects. Int. J. Neurosci. 106, 209–226.

Regan, D., 1989. Human Brain Electrophysiology: Evoked Potentials and Evoked Magnetic Fields in Science and Medicine. Elsevier, New York.

Rossion, B., Jacques, C., in preparation. The role of viewpoint and image symmetry in accounting for early face preferential responses in the human brain.

Rossion, B., Delvenne, J.F., Debatisse, D., Goffaux, V., Bruyer, R., Crommelinck, M., Guerit, J.M., 1999. Spatio-temporal localization of the face inversion effect: an event-related potentials study. Biol. Psychol. 50, 173–189.

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

Rossion, B., Gauthier, I., Goffaux, V., Tarr, M.J., Crommelinck, M., 2002. Expertise training with novel objects leads to left-lateralized facelike electrophysiological responses. Psychol. Sci. 13, 250–257.

Rossion, B., Joyce, C.A., Cottrell, G.W., Tarr, M.J., 2003. Early late-ralization and orientation tuning for face, word, and object processing in the visual cortex. NeuroImage 20, 1609–1624.

Rossion, B., Kung, C.C., Tarr, M.J., 2004. Visual expertise with nonface objects leads to competition with the early perceptual processing of faces in the human occipitotemporal cortex. Proc. Natl. Acad. Sci. U. S. A. 101, 14521–14526.

Rossion, B., Collins, D., Goffaux, V., Curran, T., 2007. Long-term expertise with artificial objects increases visual competition with early face categorization processes. J. Cogn. Neurosci. 19, 543–555.

Rousselet, G.A., Mace, M.J.M., Fabre-Thorpe, M., 2003. Is it an animal? Is it a human face? Fast processing in upright and inverted natural scenes. J. Vis. 3 (5), 440–455 (http://journalofvision.org/3/6/5/, doi:10.1167/3.6.5).

Rousselet, G.A., Mace, M.J.M., Fabre-Thorpe, M., 2004a. Animal and human faces in natural scenes: how specific to human faces is the N170 ERP component? J. Vis. 4 (2), 13–21 (http://journalofvision.org/4/1/2/, doi:10.1167/4.1.2).

Rousselet, G.A., Mace, M.J.M., Fabre-Thorpe, M., 2004b. Spatiotemporal analyses of the N170 for human faces, animal face's and objects in natural scenes. NeuroReport 15, 2607–2611.

Rousselet, G.A., Husk, J.S., Bennett, P.J., Sekuler, A.B., 2005. Spatial scaling factors explain eccentricity effects on face ERPs. J. Vis. 5 (1), 755–763 (http://journalofvision.org/5/10/1/, doi:10.1167/5.10.1).

Rousselet, G.A., Husk, J.S., Bennett, P.J., Sekuler, A.B., 2007. Single-trial EEG dynamics of object and face visual processing. NeuroImage 36, 843–862.

Sagiv, N., Bentin, S., 2001. Structural encoding of human and schematic faces: holistic and part-based processes. J. Cogn. Neurosci. 13, 937–951.

Sauseng, P., Klimesch, W., Gruber, W.R., Hanslmayr, S., Freunberger, R., Doppelmayr, M., 2007. Are event-related potential components generated by phase resetting of brain oscillations? A critical discussion. Neuroscience 146, 1435–1444.

Sayers, B.M., Beagley, H.A., 1974. Objective evaluation of auditory evoked EEG responses. Nature 251, 608–609.

Schiltz, C., Sorger, B., Caldara, R., Ahmed, F., Mayer, E., Goebel, R., Rossion, B., 2006. Impaired face discrimination in acquired prosopagnosia is associated with abnormal response to individual faces in the right middle fusiform gyrus. Cereb. Cortex 16, 574–586.

Schneider, B.L., DeLong, J.E., Busey, T.A., 2007. Added noise affects the neural correlates of upright and inverted faces differently. J. Vis. 7(4) (4), 1–24 (http://journalofvision.org/7/4/4/, doi:10.1167/7.4.4).

Schroeder, C.E., Mehta, A.D., Givre, S.J., 1998. A spatiotemporal profile of visual system activation revealed by current source density analysis in the awake macaque. Cereb. Cortex 8, 575–592.

Schroeder, C.E., Molholm, S., Lakatos, P., Ritter, W., Foxe, J.J., 2004. Human–simian correspondence in the early cortical processing of multisensory cues. Cogn. Process 5, 140–151.

Schweinberger, S.R., Pickering, E.C., Jentzsch, I., Burton, A.M., Kaufmann, J.M., 2002. Event-related brain potential evidence for a response of inferior temporal cortex to familiar face repetitions. Cogn. Brain Res. 14, 398–409.

Schweinberger, S.R., Huddy, V., Burton, A.M., 2004. N250r: a face-selective brain response to stimulus repetitions. NeuroReport 15, 1501–1505.

Sergent, J., Signoret, J.L., 1992. Varieties of functional deficits in prosopagnosia. Cereb. Cortex 2, 375–388.

Sergent, J., Ohta, S., Macdonald, B., 1992. Functional neuroanatomy of face and object processing—a positron emission tomography study. Brain 115, 15–36.

Severac-Cauquil, A.S., Edmonds, G.E., Taylor, M.J., 2000. Is the face-sensitive N170 the only ERP not affected by selective attention? NeuroReport 11, 2167–2171.

Shibata, T., Nishijo, H., Tamura, R., Miyamoto, K., Eifuku, S., Endo, S., Ono, T., 2002. Generators of visual evoked potentials for faces and eyes in the human brain as determined by dipole localization. Brain Topogr. 15, 51–63.

Sorger, B., Goebel, R., Schiltz, C., Rossion, B., 2007. Understanding the functional neuroanatomy of acquired prosopagnosia. NeuroImage 35, 836–852.

TallonBaudry, C., Bertrand, O., Delpuech, C., Pernier, J., 1996. Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. J. Neurosci. 16, 4240–4249.

Tanaka, J.W., Curran, T., 2001. A neural basis for expert object recognition. Psychol. Sci. 12, 43–47.

Tanskanen, T., Nasanen, R., Montez, T., Paallysaho, J., Hari, R., 2005. Face recognition and cortical responses show similar sensitivity to noise spatial frequency. Cereb. Cortex 15, 526–534.

Tarkiainen, A., Cornelissen, P.L., Salmelin, R., 2002. Dynamics of visual feature analysis and object-level processing in face versus letter-string perception. Brain 125, 1125–1136.

Thierry, G., Pegna, A.J., Dodds, C., Roberts, M., Basan, S., Downing, P., 2006. An event-related potential component sensitive to images of the human body. NeuroImage 32, 871–879.

Thierry, G., Martin, C.D., Downing, P., Pegna, A.J., 2007a. Controlling for interstimulus perceptual variance abolishes N170 face selectivity. Nat. Neurosci. 10, 505–511.

Thierry, G., Martin, C.D., Downing, P., Pegna, A.J., 2007b. Is the N170 sensitive to the human face or to several intertwined perceptual and conceptual factors? Nat. Neurosci. 10, 802–803.

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

VanRullen, R., 2006. On second glance: still no high-level pop-out effect for faces. Vis. Res. 46, 3017–3027.

VanRullen, R., Thorpe, S.J., 2001. The time course of visual processing: from early perception to decision-making. J. Cogn. Neurosci. 13, 454–461.

Verfaillie, K., Boutsen, L., 1995. A corpus of 714 full-colour images of depth-rotated objects. Percept. Psychophys. 57, 925–961.

Watanabe, S., Kakigi, R., Puce, A., 2003. The spatiotemporal dynamics of the face inversion effect: a magneto- and electro-encephalographic study. Neuroscience 116, 879–895.

Yeung, N., Bogacz, R., Holroyd, C.B., Cohen, J.D., 2004. Detection of synchronized oscillations in the electroencephalogram: an evaluation of methods. Psychophysiology 41, 822–832.

Yin, R.K., 1969. Looking at upside–down faces. J. Exp. Psychol. Hum. Percept. Perform. 81, 141–145.

Zion-Golumbic, E., Bentin, S., 2007. Dissociated neural mechanisms for face detection and configural encoding: evidence from N170 and induced gamma-band oscillation effects. Cereb. Cortex 17, 1741–1749.