DOCTOR OF PHILOSOPHY

What is special about faces? Examining face-categorisation with event-related potential measures

Dering, Benjamin

Award date:
2012

Awarding institution:
Bangor University

What is special about faces? Examining face-categorisation with event-related potential measures

Doctoral dissertation presented by:

Benjamin Robert Dering

Supervisor:

Prof. Guillaume Thierry

Second Supervisor:

Dr. Stephan Boehm

Chair:

Dr. Paloma Mari-Beffa

What is special about faces? Examining face categorisation with event-related potential measures.

Contents

Chapter 1

Thesis overview and general introduction

Overview of thesis:

Multiple independent lines of research have suggested that faces are a special class of stimulus. In the last 15 years, neuroimaging studies have shown greater activation to faces than to any other stimulus category in specific areas of cortex, leading to the idea that a portion of the fusiform gyrus, also known as the fusiform face are (FFA), is face-selective (Kanwisher, McDermott, & Chun, 1997). While findings from neuroimaging, behavioural, and lesion studies support the idea of a specialised visual system for faces, it is still debated whether face sensitivity arises from either an inherent face modular network or a general processing network manifesting perceptual expertise. A modular network is an abstract cognitive concept representing functions of the brain that require rapid, automatic cognitive processing. Modules are argued to be domain specific and information encapsulated such that they do not need to interact with other cognitive processes to function. In contrast to face modularity, the expertise account of face processing argues that faces recruit domain general processing mechanisms, which are not unique to faces, but finely tuned from extensive perceptual experience. In other words, the expertise account considers faces as stimuli for which almost everyone is a skilled expert.

Attempts to make progress in the debate opposing domain specific (modular) vs. domain general (expertise-based) processing has led to investigations into neurophysiological indices of face processing. Alongside the vast behavioural literature portraying the human face as a unique and special visual stimulus, electrophysiological studies have focused on a negative polarity component from the N1 family and peaking at around 170 ms, the N170. The N170 is maximal over parietal-occipital electrode sites, and widely acknowledged as largest in amplitude to faces (Bentin, Allison, Puce, Perez, & McCarthy, 1996). Since the seminal study by Bentin et al. (1996), it has been claimed repeatedly that no visual stimulus other than faces produces negativities as pronounced in the N1 range (Itier & Taylor, 2004a). So robust is this finding that the N170 face effect has been replicated and championed to the point where it is no longer considered a hypothetical effect but rather an established fact (Eimer, 2011; Rossion & Jacques, 2008).

Like fMRI, electrophysiology cannot directly elucidate the debate concerning modularity vs. expertise-based processing, since face-selectivity, observable in ERPs as an amplitude increase for faces in the N1 range, can be predicted by both theoretical standpoints. In contrast to the majority of the ERP literature, there are however, instances where face-selectivity in the N170 range was not found, particularly in the case of studies comparing full-front views of objects such as cars and butterflies (Rossion, et al., 2000; Schweinberger, Huddy, & Burton, 2004; Thierry, Martin, Downing, & Pegna, 2007a). For instance, Thierry et al. (2007a) showed that inter-stimulus variability within an object class, a factor mixing physical and perceptual variance, modulates the amplitude of the N170 component. When comparing faces with other categories of object, previous studies have often used faces presented full front, in an upright orientation, whereas contrasting object classes have often been variable in size, background, orientation, viewpoint, etc. This may have lead to imbalanced experimental comparisons artificially increasing the N170 elicited by faces because of the low inter-stimulus variance usually inbuilt for faces in the design. Thierry et al. (2007a) compared full front views of faces with full front views of cars or butterflies and found no significant mean amplitude differences between conditions in the N170 range. Furthermore, they reported category-sensitivity unaffected by inter-stimulus variance 70 ms earlier, in the range of the P1 (~100 ms after stimulus onset). It is noteworthy that P1 face-sensitivity has been largely overlooked in previous research, despite some reports, which have highlighted such potential sensitivity (Herrmann, Ehlis, Ellgring, & Fallgatter, 2005; Herrmann, Ehlis, Muehlberger, & Fallgatter, 2005; Linkenkaer-Hansen, et al., 1998). In sum, Thierry et al. (2007a) questioned the validity of object categorisation experiments, which used stimuli varying in many more ways than object category, particularly in terms of low-level perceptual features.

Thierry et al.'s results (2007a) have been staunchly refuted (see (Bentin, et al., 2007; but also Thierry, Martin, Downing, & Pegna, 2007b). In fact, Rossion & Jacques (2008) dedicated a review article to the dismissing of the arguments put forward by Thierry et al. (2007a). In this publication, they present new data, using a design very similar to that of Thierry et al. (2007a) but displaying face-selectivity in the N170 range. The conflicting findings of Rossion and Jacques, (2008) and Thierry et al. (2007a), and the heavy criticism by Rossion & Jacques (2008) of Thierry et al.'s conclusions, have created some confusion within the field, questioning the established view that N170 reflects visual object categorisation.

This thesis is concerned with the further characterization of stages in face processing as indexed by ERPs. Specifically, I question the point in time at which ERP waveforms can detect the first observable differences between faces and other objects, and whether these differences are indicative of a specialised process dedicated to faces. I present a series of ERP experiments explicitly testing the category sensitivity of early ERP components, namely the P1 and the N170, since their functional significance remains poorly understood. A subtheme of the thesis is to determine whether differences in ERP component amplitude constitute a reliable measure of face (and –more generally– object category) sensitivity, and if so, whether these differences are attributable to early object categorisation or higher level processes such as individual object recognition / identification. [1]

More specifically, in the present work, the aim is to address the following questions: (1) Can Thierry et al. (2007a) be replicated, and does the task involved interact with the commonly accepted N170 category-selectivity? (2) Does inter-stimulus perceptual variance affect/interact with the N170 face inversion effect? (3) Do any other perceptual parameters affect P1 & N170 amplitude? For instance, does cropping faces out of heads modulate P1/N170? (4) If one creates face-car hybrid using morphing algorithms, do the P1 and N170 reflect the amount of face information present in the stimulus? (5) Can expertise with complex visual stimuli entirely account for the N170 inversion effect?

At this point, it is important to make a distinction between selectivity or specificity on the one hand and sensitivity on the other. To make a genuine

---

[1] For the purpose of clarity throughout this thesis, I will refer to object categorisation when discussing the distinction between different categories of objects and to recognition or identification when discussing the extraction of higher-level properties such as ethnic origin, emotional expression, intention, age, gender, and even familiarity or identity.

claim regarding category selectivity/specificity[2], one would have to test objects from every single existing conceptual category in comparison to faces (in the present case). However, an ERP component can be sensitive to a particular category of objects when its amplitude and/or latency is modulated by categorical changes, without a need for exhaustively testing all existing categories, as if this was humanly possible.

---

[2] Throughout this thesis the terms selective/specific will be used when referring to previous research making claims in support of N170 face selectivity, whereas the term sensitivity will be used to refer to the present results and any conclusions drawn from them.

Introduction

One fundamental aspect of human cognition is the distribution of sensory information to neural areas that implement processing, best described in the visual system. Upon reaching primary visual cortex (V1), visual information is commonly thought to split into two streams of information processing, the dorsal stream and the ventral stream (Milner & Goodale, 1995). These streams project to different brain regions, processing different aspects of vision like colour (V4 – ventral stream) and motion (MT – dorsal stream; Zeki, et al., 1991). The repartition of visual information to distinct brain regions, like V4 and MT, is symbolic of the idea that not only human cognition, but the brain itself can be anatomically segregated into functional modules. A module, the base unit of the modularity hypothesis, is an independent, encapsulated information processing system, which does not receive interference from other processing networks (Fodor, 1983). This view can be extended to include aspects of cognition taking place at a higher level than low-level differentiations in visual cortex. Face processing is prototypical of the modularity view; it is distinguished from other forms of visual object processing by theoretical models (Bruce & Young, 1986), reaction time data, lesion studies of prosopagnosic patients (Farah, 1996; Farah, Levinson, & Klein, 1995; Farah, Wilson, Drain, & Tanaka, 1995), single cell recordings (Oram & Perrett, 1992; Perrett, Oram, & Ashbridge, 1998; Perrett, et al., 1984), and neuroimaging (Kanwisher, 1997).

Faces are special because they relay a wealth of personal information like no other object; gender, age, race, emotion and attractiveness can be evaluated in an instant. Before recognition of any object in a visual scene, a face needs to be defined from the background context first. At this lowest level, light intensity changes (luminosity) and edge detection (driven by contrast) are thought to be the earliest processes implemented by the visual system. Marr (1982) describes this stage of early vision as the primal sketch. Further computations by the visual system take place at a higher cognitive level, utilising binocular cues, texture information, and shape to build an accurate viewpoint-dependent representation of the object. As a result, at the lowest level of vision, faces and objects may be treated similarly, due to the generic nature of distinguishing contrast, luminance, spatial frequencies, and edges (see Biederman, 1987; for an extension upon Marr's work). Processing of higher-level attributes such as gender, race, emotion and attractiveness of a face, are likely to call upon far more detailed analysis of object properties and content such as textures, colours, configuration. In the work reported in this thesis, a main object of focus is that of categorisation, i.e., the moment in time during visual object processing when the first observable differences between different categories of object arise using event related potentials (ERPs). But first, I will discuss the cases for and against a modular basis for face processing and the specialised cognitive processes that appear to distinguish faces from other visual objects.

Theories of face processing:

The modularity hypothesis has been used as a framework for theories of face processing. The most influential of these is the Bruce and Young's (1986) model of face recognition. In this model, after an initial stage of structural encoding, the presentation of a face leads to the systematic activation of nodes representing higher-level face-recognition (see fig.1 for an overview).

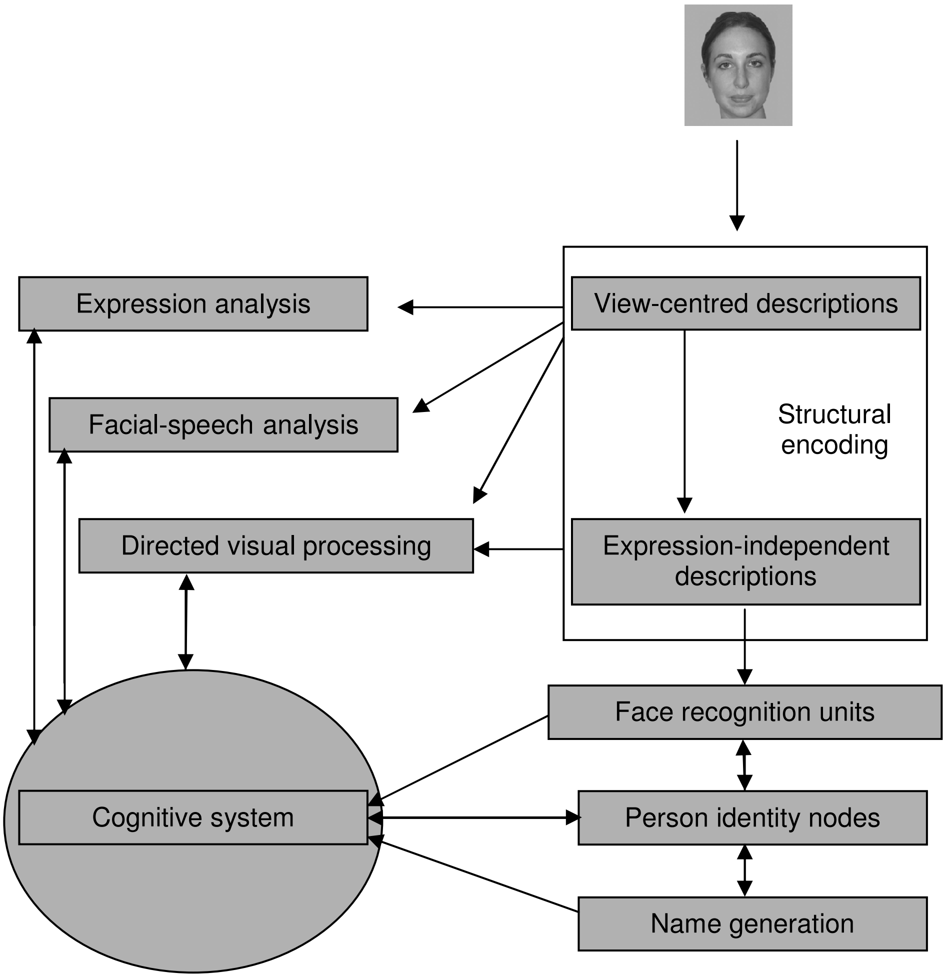

Figure 1 | **Bruce and Young's (1986) model of face recognition. Upon viewing a face, the image enters a structural encoding phase based upon orientation and configural information. At the next level, familiar face identification follows the pathway of face-recognition units, while expression analysis, facial speech analysis, and directed visual processing are more important for processing unfamiliar faces. While still essential, the use of this model's concepts in recent neurophysiological work has some shortcomings. For example, in the model, top-down knowledge, represented by the cognitive system, does not interact with structural encoding. Furthermore, It is**

---

unclear from this model how unfamiliar face identification processes, e.g., expression analysis, interact with familiar face recognition.

Structural encoding refers to the production of orientation specific (e.g., profile view or full-front) and expression nonspecific (i.e., independent of emotional state) representations of faces, and is considered essentially modular in nature. Higher levels of processing in Bruce & Young's (1986) model cannot interact with this critical stage. While the model is important for conceptualising late face recognition processes such as name generation, it does not provide any insight into the initial perceptual processes underlying object categorisation. For instance, it is unclear if structural encoding encompasses face categorisation, since this model was conceived to deal with faces exclusively. The authors themselves concede that face categorisation probably originates from object processing mechanisms, but they speculate that categorisation is part of structural encoding by default. Moreover, in the model, structural encoding incorporates higher level face processes such as face identification before information pathways split between processing of familiar and unfamiliar faces.

Other theories of face processing have attempted to characterise the particular cognitive processes involved in discriminating faces from other objects. For instance, Farah, Wilson, Drain, & Tanaka, (1998) suggested that faces are processed holistically, such that individual feature differentiation is not pivotal in the analysis of the face as a whole. In comparison, objects are processed by an analysis of their configural relations between parts, i.e., the relative spatial location of their constituent parts. According to Farah, Levinson, et al., (1995), this distinction in processing mechanisms is attributable to a face-specific module allowing rapid holistic processing of face information. Support for such modular processing comes from studies of individuals suffering from acquired prosopagnosia, a cognitive deficit occurring after stroke or lesion in which the ability to recognise faces is severely impaired. In contrast, object processing is relatively intact (Farah, 1996; Farah, Wilson, et al., 1995), leading to a single dissociation between face and object processing in prosopagnosic patients. However, this single dissociation has been challenged, because it is unclear whether prosopagnosia maybe an extreme form of object agnosia, affecting the recognition of face-objects to a far greater extent than that of other object categories (Damasio, Damasio, & Van Hoesen, 1982; but see De Renzi, 2000). It must also be noted that most prosopagnosic individuals have no difficulty in categorising faces as faces, and even discriminating sex or race information, but their impairment concerns individual face recognition.

Whilst data from prosopagnosic patients supports the concept of modularity in the face processing system, since a double dissociation has been suggested between pure prosopagnosia and pure visual agnosia (sparing faces, Moscovitch, 1997), inverting faces provides the most compelling evidence that two distinct processing mechanisms operate for faces and other objects. If faces are displayed upside down, recognition is disproportionately less accurate as compared to the case of inverted objects. This effect is classically described as the face inversion effect (Yin, 1969; Fig. 2a). Farah, Tanaka, & Drain (1995), for instance, showed that inverting faces disrupts a holistic mechanism specialised for upright views of faces in groups of participants trained to perceive patterns or faces either as whole images or to focus on their inner parts. It is only when participants relied on holistic processing that inverted stimuli delayed reaction times and decreased accuracy. This is consistent with the lack of a face inversion effect in prosopagnosic patients, suggesting that they rely on an analysis by parts rather than a holistic one (Farah et al. 1995).

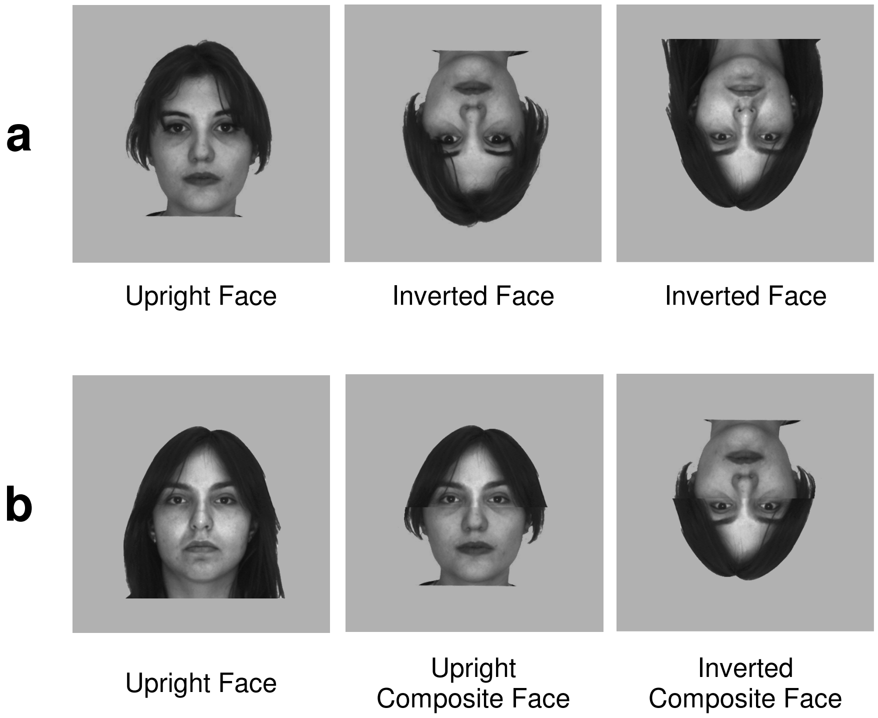

Figure 2 | **Stimulus manipulations used to test holistic processing. (a) Face inversion. The face inversion effect is characterised by an increase in errors and a delay in reaction time greater for faces than other objects. Typically this detriment in performance is attributed to a switch in processing mechanisms, from a holistic analysis to an analysis of the spatial locations of constituent parts, when looking at faces. (b) The composite face effect. The top half of the upright face and the upright composite are the same. It is more difficult to recognise two distinct images when the**

composite is upright than when inverted. This is taken as evidence for holistic processing of faces when presented in an upright orientation.

In contrast with Farah’s (1998) modular view of face processing, Diamond & Carey (1986) have proposed that visual objects are recognised based on relationships (a) between the object and its context (first order relations) and (b) between internal constituent parts of objects (second order relations). First order relations apply to classes of objects that have no specific shared configuration, such that the relation between parts can differ, e.g., a landscape and objects within it. Second order relations apply to structured object classes that have a generic configuration, e.g., like faces that contain always two eyes, a nose and mouth. Diamond & Carey (1986) theorised that second order configural relations not only predominantly drive individual face processing, but must be influential in processing individual objects from other classes. Face processing is therefore not implemented by a functionally independent module, but rather belongs to domain general visual recognition.

However, in cases where the observer is inexperienced with a particular type of object, second order configural relations are far less informative. Diamond & Carey (1986) demonstrated that extensive perceptual experience with a particular object class is required for second order configural information to play a significant role. They presented pictures of dogs in two orientations (upright or inverted) to dog experts and participants with little experience of dogs. Only the experts showed a substantial increase in reaction time and decrease in accuracy for inverted dogs as compared to upright dogs, an effect similar in magnitude as that seen for face inversion (Rossion & Curran, 2010). However, many studies have failed to find effects of expertise in object recognition, partly because perceptual experience with common objects is very difficult to measure and control, since expertise is essentially based on unreliable self-report measures (Bruyer & Crispeels, 1992; Busey & Vanderkolk, 2005; Gauthier, Skudlarski, Gore, & Anderson, 2000; Gauthier, Williams, Tarr, & Tanaka, 1998; Robbins & McKone, 2007). Gauthier & Tarr (1997) controlled for participant's perceptual experience using a novel class of object (Greebles) and trained participants to be experts with them. When tested using upright and inverted orientations, inversion effects representative of second order configural processing, were not found. The diverse findings of expertise studies indicate that perceptual expertise, whether measured in years or completion of a training paradigm, are perhaps not enough to enforce a switch in processing mechanisms from part-based first order configural relations to global second order configural relations, strengthening the case for a modular face network.

Maurer, Grand, & Mondloch (2002) attempted to merge the two main theoretical perspectives described above. Maurer et al. (2002) refer to (1) first order relations as the set configuration of internal face features i.e., eyes above nose above mouth (fundamental for the categorisation of a face as a face); (2) holistic processing as the perception of a face as a whole; and (3) second order relations as the distance between internal face features (fundamental for individual recognition). These stages in configural processing are hypothesised to occur in this order. Thus, this model is in fact a hybrid of the models put forward by Farah et al. and Diamond and Carey. Unlike Farah et al. (1995), however, Maurer et al. (2002) do not debate whether holistic processing is inherently modular or develops through perceptual experience.

In summary, the above mainstream theories provide a skeleton model for elementary stages of face processing, which may or may not have generic value for visual object recognition. It is noteworthy that these theoretical conceptualisations have been built essentially on experimental data acquired idiosyncratically using specific object categories (e.g., faces, dogs, birds, cars, greebles, considered separately). Indeed, none of the above theories present a complete picture of face recognition processes, starting with what differentiates faces from other objects. For example, Bruce and Young's model does not discuss how object categorisation can lead to distinctions between faces and other objects, but is rather an account of later stages in face recognition beyond initial object categorisation. Is it that the human face is such an important and biological relevant stimulus that it must be dealt with in a specialised, isolated way? In the same vein, Farah et al's two-process model is an oversimplification of general object recognition, which excludes processes such as edge detection (driven by contrast distinctions, Marr, 1982), even though it must be fundamentally involved in object categorisation. With the rise of neuroscience methods in the 1990s, a whole branch of research has developed to substantiate theoretical models of visual objects processing, and more specifically models of face processing.

Selective review of electrophysiological and neuroimaging research

The broadening field of cognitive neuroscience employs neuroimaging methods such as functional magnetic resonance imaging (fMRI), single cell recording, electroencephalography (EEG) and magnetoencephalography (MEG) to further understand how faces are processed by the human brain. Studies using fMRI have found a greater blood oxygen level-dependent (BOLD) response to faces than many other objects in the fusiform gyrus, which has since been labelled the fusiform face area (FFA; Kanwisher, et al., 1997; Kanwisher & Yovel, 2006). While there are conflicting viewpoints regarding the face-specificity of the FFA (e.g., Bukach, Gauthier, & Tarr, 2006), multiple other regions of cortex also display face sensitivity, e.g., the occipital face area and the posterior superior temporal sulcus (Taylor & Downing, 2011). Instead of the whole of face recognition being implemented in a single area of the visual association cortex such as the FFA, it is likely that many other areas of cortex, in combination with FFA, form a modular face processing network. Gschwind, Pourtois, Schwartz, Van De Ville, & Vuilleumier (2011) subjected data from a face localiser task to diffusion tensor imaging (DTI) analysis, a probabilistic method of determining pattern direction to effectively map white-matter fibre tracts. The authors showed strong white matter connections between FFA and OFA. In contrast, no significant connections between STS and FFA were found. However, this might be due to the fact that the face localiser task used static images of faces, while it has been shown that the STS is more sensitive to dynamic facial movements than static images (Pitcher, Dilks, Saxe, Triantafyllou, & Kanwisher, 2011). Furthermore, in Gschwind et al. (2011), amygdala activity bypassed FFA and was connected with early visual areas in occipital cortex, suggesting that the

FFA may not be involved in the emotional processing of faces. Overall, these results, along with work by Gobbini & Haxby (2007) show that if face processing is to be considered a modular process, it is instantiated by a network of regions distributed throughout the brain.

Proponents of the expertise account of face-processing argue that it is still undetermined whether face-sensitive areas of cortex really reflect processing domain-specific to faces or may equally correspond to a domain-general visual system. Indeed, there is evidence for modulation of right FFA activation by expertise with non-face stimuli, although this is commonly attributed to mere face-likeness effects (Bilalic, Turella, Campitelli, Erb, & Grodd, 2011; Gauthier, et al., 2000; Xu, Liu, & Kanwisher, 2005). To the author's knowledge, evidence in support of the modularity view of face processing using fMRI has exclusively used face localiser tasks to identify cortical regions of interest, a task comparing faces to a set of scrambled images, or objects from mixed categories.

Single cell recording studies have identified populations of cells selective to face stimuli (Oram & Perrett, 1992; Perrett, et al., 1998; Perrett, et al., 1984). In strong support of the modularity view of face processing, Tsao, Freiwald, Tootell, & Livingstone (2006) showed that cortical regions of face sensitivity in Rhesus Macaques, identified with fMRI, consist almost entirely of cells with an increased sensitivity to faces. Of the cells recorded from in the STS (320 across 2 Macaques), 97% had a stronger firing rate to faces than to other objects. The region of cells tested in both macaques was believed to be topographically homologous to the FFA in humans (Tsao et al., 2006). The authors concluded that a face-selective region of cells in the brain would be needed to support specialised holistic processes attributed to a modular face processing network.

Unfortunately, no current imaging technique effectively maps distributed neuronal functioning. In the case of fMRI for instance, the quantification of oxygenation in red blood cells that is indexed by the BOLD response is fundamentally an indirect measure of neuron activity. Therefore, fMRI is best used in conjunction with other techniques, such as single cell recording, and EEG, etc. in order to cross-validate theoretical accounts put forward separately. Intracranial recordings give an accurate description of neuron firing. They have both high temporal resolution and precise locations in space, but they are also highly invasive, and lesions can occur upon insertion. Because of the invasive nature of this technique, instances for use with human subjects are few, and limited to neurologically abnormal patients. Also, generation of a scalp potential measurable with EEG requires synchronous firing of roughly 6×10^7 neurons over a 6 cm^2 area of cortical tissue (Ebersole, 1997). Therefore, findings from intracranial recordings are not directly comparable to EEG or even fMRI data, since the number of cells recorded from is miniscule in comparison to the ongoing synaptic activity in the area.

EEG measures the global change in voltage across the scalp non-invasively. The scalp potentials recorded to stimulus presentation are averaged to form an event-related potential (ERP) wave; the resulting wave summarises the average brain activity elicited by a particular cognitive event

(Fig. 3). Unlike fMRI and single cell recording, ERPs have poor spatial resolution. Whilst spatial location can be determined with accuracy in the order of millimetres, it is not possible to determine the source of neural generators since the summed activity of a potentially infinite number of neural generators can produce the same scalp topography. The search for a solution starting from the surface topography is known as the inverse problem. Hence, source localisation procedures, while reliable and accurate, might not provide valid descriptions of underlying neural generators. However, the advantage of EEG over fMRI and single cell recording lies in the temporal domain, where recordings can be made with sub-millisecond precision. Single cell studies can record with a high temporal resolution also, yet do not give information over global changes in brain activity, while measurement of the fMRI BOLD response is in the order several hundred milliseconds. Capitalising on the temporal resolution of ERPs, studies of face processing have highlighted a negative modulation for faces 170 ms after stimulus onset, the N170, first described by Bentin et al. (1996).

The N170 component of ERPs

The N170 is a negative deflection of the ERP signal over posterior-occipital regions of the scalp, supposedly greater in amplitude to faces than any other object category (Fig. 3). Since the publication of Bentin et al. (1996), many reports explicitly refer to the N170 as a face-selective component (e.g., (Bentin, et al., 1996; Blau, Maurer, Tottenham, & McCandliss, 2007; Carmel & Bentin, 2002; Crist, Wu, Karp, & Woldorff, 2007; Eimer, 2000b, , 2011; Eimer, Kiss, & Nicholas, 2010; Itier & Taylor, 2002; Mohamed, Neumann, & Schweinberger, 2009; Rossion & Jacques, 2008; Sadeh, Zhdanov, Podlipsky, Hendler, & Yovel, 2008).

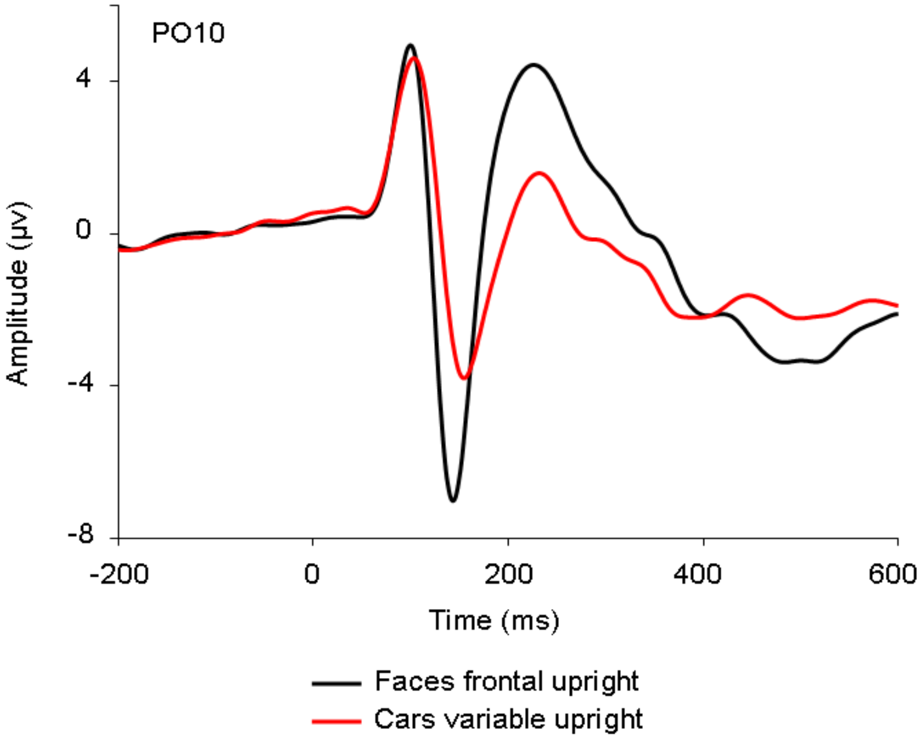

Figure 3 | **An example of an ERP waveform at electrode site PO10. The first negative wave (N1) is often termed the N170. Here, faces (black) show a greater increase in amplitude at the N170 in comparison to contrasting object categories, in this case cars (grey). Data adapted from Boehm, Dering & Thierry (2011), *Neuropsychologia*.**

The N170 has been strongly linked to the idea of structural encoding, as proposed in the model of Bruce & Young (1986), representing the initial stages of face categorisation and identification, with module-like characteristics, i.e., it is information encapsulated and cognitively impenetrable (Pylyshyn, 1999). Eimer et al. (2000) suggested that the N170 component is related to late stages of structural encoding of faces, namely face identification, given that when internal features of faces are masked in profile views, or rear views, the N170 is attenuated in comparison to that elicited by full-front views of upright faces. Nevertheless, the N170 is considered larger in amplitude to faces than any other object, regardless of their orientation. Even though the evidence for face-sensitive processing occurring in the N170 range is overwhelming, the exact nature of the cognitive mechanisms indexed by the N170 is still debated. As noted earlier (see Fig.1.), the structural encoding phase of Bruce & Young's (1986) model does not interact with higher level face identification processes, such as emotion recognition or name generation, and can be thought of as an initial gating mechanism for later face processing stages. However, many publications have reported evidence for top-down modulations of the N170 component by familiarity (e.g., Eimer, et al., 2010; Vizioli, Rousselet, & Caldara, 2010) and by emotional expression (e.g., Dubal, Foucher, Jouvent, & Nadel, 2011). Therefore, strictly speaking according to the Bruce & Young (1986) model, the N170 cannot represent processing of a structural encoding phase of faces.

Further support for the sensitivity of the N170 to face processing comes from the observation that it is increased in amplitude, and delayed in latency, for faces presented upside-down as compared to faces presented upright (Eimer, 2000a; Rossion, et al., 1999). In addition, it is particularly striking that the inversion effect on N170 amplitude and latency, is often not found for any other object than faces (Itier & Taylor, 2004a). The absence of N170 inversion effects for other objects, which has led to the common understanding that it is only observed for faces, is reminiscent of the behavioural inversion effect first described by Yin (1969), and contradicts the idea that face processing is a matter of expertise only, as proposed by Diamond & Carey (1986).

Possible reasons why stimulus inversion is not often reported with other objects than faces are that:

(1) Critically, face inversion is a test of identification (i.e., whose face is this?) whereas other objects are generally recognised at a more generic level (i.e., what object is this?). Indeed, patients and control participants are rarely if ever asked to determine whether a face presented upside-down is a face or not. It is likely that performance on such a task would be extremely high and on a par with inverted objects from other categories. It remains unclear whether participants asked to identify unique complex objects presented in inverted orientation would display an inversion effect comparable to that observed for face identification.

(2) Object categories compared to faces often lack the complexity and sophistication of face stimuli. More generally, faces are one of the most complex stimulus category encountered in everyday life. Most other objects that have been compared to faces, such as houses, butterflies, watch-dials, have excessively simpler structure and second order relations in the sense of Diamond and Carey (1986).

(3) Participant’s perceptual experience with other object categories is almost impossible to control, as already alluded to. Indeed, perceptual experience with stimuli is likely to affect the strength of the N170 inversion effect. It is possible that high levels of expertise with a particular object class,

leads observers to holistic processing strategies of the kind used for faces. Consistent with this hypothesis, visual expertise has been shown to modulate hemodynamic responses in the FFA (Gauthier, et al., 2000; Gauthier, Tarr, Anderson, Skudlarski, & Gore, 1999) as well as N170 amplitude (Tanaka & Curran, 2001; Tanaka, Luu, Weisbrod, & Kiefer, 1999).

In addition to the shortcomings affecting inversion effects listed above, it still remains entirely unclear to what extent the N170 may or may not reflect modular domain-specific processing or expertise-based processing. Unfortunately, N170 studies in patients with prosopagnosia have not helped much in progressing this debate, since some studies have reported an extinction of the N170 face effect (e.g., (Eimer & McCarthy, 1999), while others found no correlation between N170 modulation and prosopagnosic symptoms (Harris, Duchaine, & Nakayama, 2005).

N170 or P1?

Thierry et al. (2007a) reported results from a series of experiments manipulating inter-stimulus perceptual[3] variability (ISPV), in which faces did not elicit greater N170 amplitude than contrasting object categories (i.e., cars and butterflies). Their argument revolves around the observation that faces are commonly presented in one viewpoint, orientation, size, and position, whereas images of other objects used in previous studies were often more

---

[3] Note here that Thierry et al. originally used the word “perceptual” to describe inter-stimulus variance in their 2007a study. Even though Rossion and colleagues (e.g., Rossion and Jacques, 2008) swapped the term for “physical” in an attempt to shift the focus of attention onto the physical disparity of stimuli presented within an experiment, this factor is not the factor that Thierry et al. (2007a) intended to manipulate in their paper as clearly stated in the legend of their Fig. 2.

variable regarding all of these parameters. It must be noted that a number of these studies have used such comparisons to make claims regarding the category selectivity of the N170 component (e.g., Botzel, Schulze, & Stodieck, 1995; Carmel & Bentin, 2002; Itier, Latinus, & Taylor, 2006; Itier & Taylor, 2004b; Philiastides & Sajda, 2006; Rossion, et al., 2000; Rousselet, Husk, Bennett, & Sekuler, 2005; Rousselet, Mace, & Fabre-Thorpe, 2004). In experiment 1 of Thierry et al. (2007a), ISPV (low / high) and object category (face / car) were tested in a two-by-two design. Surprisingly, no differences between face and car conditions were found for low ISPV stimuli. Furthermore, this result was essentially replicated when profile views of faces where compared to profiles views of butterflies, an object category sharing little or no features with faces. Importantly, Thierry et al. (2007a) also reported that the earlier P1 component (peaking at around 100 ms after stimulus onset) displayed the sensitivity expected of the N170 to ISPV and categories if it had been purely face-selective, namely no sensitivity to ISPV and significant increase in amplitude for faces as compared to cars or butterflies.

The concept of perceptual variance between conditions must be distinguished from that of physical variance since two images may be physically identical in terms of contrast, luminance, spatial frequencies, viewpoint, size, orientation and position, yet equalising them in terms of all these physical properties would result in evident perceptual distortion. Reciprocally, physical differences between two pictures of faces (or other objects) may well exist in the absence of any significant perceptual shift. For instance, in the extreme, the very same face image, perceptually identified as face A, may be physically altered by overlapping visual noise, or changing contrast and luminance, whilst remaining unambiguously face A (Fig. 4).

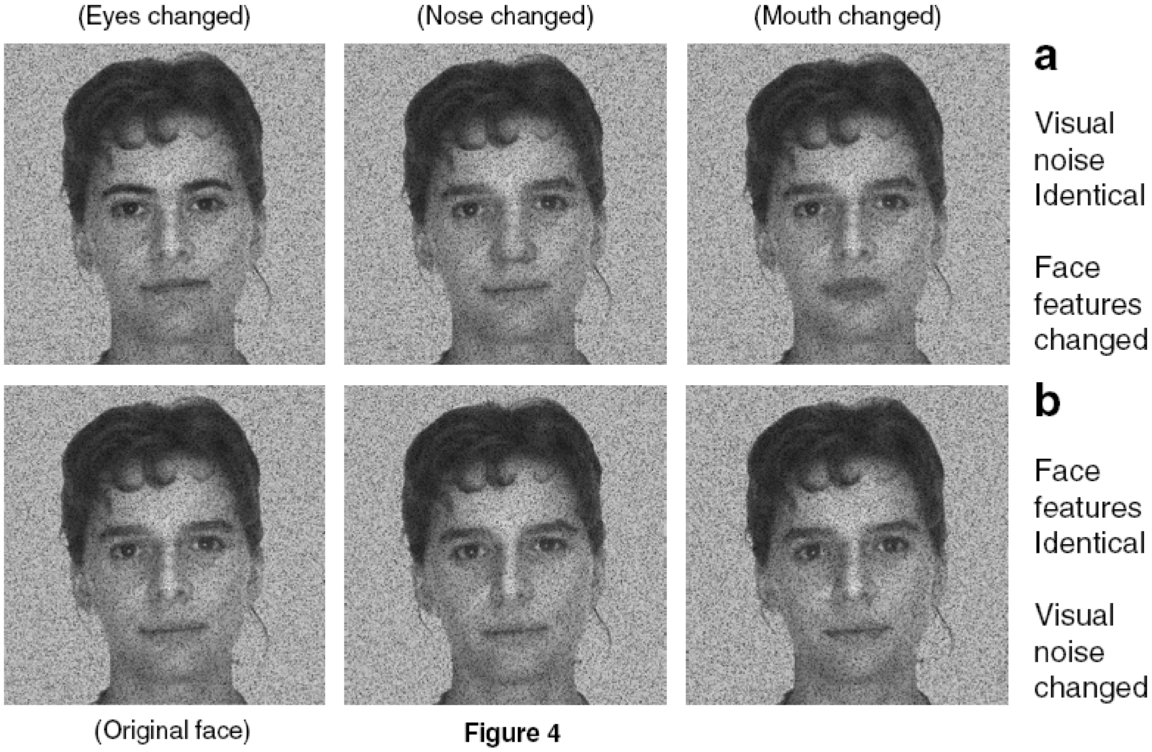

Figure 4

Figure 4 | **Demonstration of the dissociation between physical and perceptual variance. Physical variance is manipulated minimally in the upper row (a) since visual noise is identical between the different faces and only one face part has been changed (eyes, nose or mouth) and maximally in the lower row (b) since the visual noise in each picture changes, even though the face image remains the same. Arguably the faces on the top row are perceptually different from one another (they look like they belong to different individuals), whereas the faces on the bottom row are perceptually identical (the face is the same). In other words, row a displays high physical similarity and perceptual diversity whilst row b shows low physical disparity and high perceptual similarity: the two concepts dissociate entirely.**

Even though Thierry et al. (2007a) were not the first authors to compare full-front images of faces and cars and fail to find differences in N170 amplitude between these conditions (Rossion et al., 2000, Schweinberger, Huddy & Burton, 2004), they were the first to overtly manipulate stimulus variance and establish N170 sensitivity to ISPV. It must be noted category-sensitivity in P1 range has also been observed before (Herrmann, Ehlis, Ellgring, et al., 2005; Herrmann, Ehlis, Muehlberger, et al., 2005; Linkenkaer-Hansen, et al., 1998), although many studies have simply failed to analyse and report P1 effects (e.g., (Horovitz, Rossion, Skudlarski, & Gore, 2004).

Why should the P1 be category sensitive? The P1 is considered an index of low-level visual properties such as contrast, luminance and spatial frequency, and is known to be modulated by spatial attention tasks (Hopf & Mangun, 2000). It is surprising therefore that Thierry et al. (2007a) entirely failed to see any difference in the P1 range between low ISPV and High ISPV conditions, which nevertheless, differed considerably in terms of low-level visual properties. At the end of the day, it is an entirely hypothetical standpoint whether P1 indexes solely low-level visual processing or whether it is also an index of higher-level cognitive integration such as category differentiation.

The publication of Thierry et al.’s (2007a) study has initiated a controversy, which has taken the form of an exchange of commentaries in Nature Neuroscience (Bentin et al. 2007, Thierry et al., 2007b), and culminating in a “review” article published in Neuroimage (Rossion & Jacques, 2008). In the latter article, Rossion & Jacques (2008) reported results of a study in preparation (methods unpublished) in which the N170 amplitude was significantly larger for low variance full front faces as compared to low variance full front cars. This data stand directly in conflict with those of Thierry et al. (2007a) and this issue is dealt with in Chapter 4 of this thesis.

Overview of the questions addressed in this thesis

In Chapter 2, we report a replication of Thierry et al. (2007a) in a different group of participants, using two different tasks (categorisation vs. identification of immediate repetition –as in Thierry et al., 2007a) to test whether the level of detail in image analysis required would affect ERP modulations in the P1 and N170 range. First, replication is fundamental when results are controversial and when one of the main findings is an absence of effect. Second, testing potential task effects was anticipated to shed light on the susceptibility of the mechanisms indexed by P1 and N170 to top-down cognitive control.

In chapter 3, we test whether inversion effects on faces and cars are similar to those observed previously when perceptual variance is manipulated. If Thierry et al. (2007a) reported a limitation of all previous studies on face processing with ERPs, then it is possible that spurious differences in variability between experimental conditions may have affected inversion effects observed for objects other than faces. To tease apart the effects of stimulus category, variability, and inversion, it is therefore necessary to conduct a 2 x 2 x 2 factorial design.

In Chapter 4, we attempted to reconcile the results (in preparation) of Rossion & Jacques (2008) with those of both Thierry et al. (2007a) and Dering et al. (2009) concerning the presence or absence of an N170 modulation for low ISPV faces and cars. Recall that Rossion & Jacques (2008) found a significant difference between full front faces and cars, but did so using cropped images of faces (without hair, ears or neck) and unaltered cars, and stimuli repeated 6 times per condition. In three studies, we tested the effect of stimulus repetition, cropping, and cross-category morphing, while controlling for low-level visual differences between conditions. Cross-category morphing mixed faces and cars in various extents, to test P1 and N170 category sensitivity on a stimulus continuum involving more subtle differences than the direct contrast between faces and cars.

In chapter 5, we tested the potential existence of an inversion effect for complex stimuli requiring expertise that are different from faces: written words. Indeed, any stimulus eliciting N170 amplitude increases with inversion other than faces would provide additional evidence that the N170 is not face-selective but rather indexes expertise with other stimuli. To achieve this, we compared two groups of participants with differing experience of English words: British natives and Asian natives (Chinese or Japanese), who acquired the alphabetical script after the age of 11.

Chapter 2

Is the N170 peak of visual event-related brain potentials car-selective?

This chapter is published as: Dering, B., Martin, C. D., & Thierry, G. (2009). Is the N170 peak of visual event-related brain potentials car-selective? *Neuroreport*, 20(10), pp. 902 – 906, doi: 10.1097/WNR.0b013e328327201d

Abstract

The N170 is a peak of event-related brain potentials commonly acknowledged to be larger in amplitude for face stimuli compared to any other visual object. Recently, the face-selectivity of the N170 has been challenged based on the observation of similar N170 amplitude to faces and cars presented full front. Here, we measured the N170 elicited by these same stimulus categories using a one-back memory and categorisation task. We found that N170 mean amplitude was significantly larger for *cars than faces* at electrode sites considered "optimal" for measuring N170 face-selectivity in the absence of task effects. Furthermore, we found evidence for category-selectivity and task-dependence in the P1 range. These results support the idea that N170 face-selectivity is formally questionable.

Introduction

The N170 event-related potential (ERP) component has been reported as face-selective in a number of publications over the past decade (Bentin, Allison, Puce, Perez, & McCarthy, 1996; Carmel & Bentin, 2002). N170 amplitude is classically reported more negative when elicited by faces than objects from any other category. Since this effect is considered a robust finding, it has important implications for theories of visual cognition. One interpretation is that the N170 component reflects category-specific processing of faces (Carmel & Bentin, 2002). This interpretation is tempting because of insights from studies of patients with focal brain lesions, on the one hand, and neuroimaging studies, on the other, which point to a ‘special’ status for face stimuli. For instance, studies of prosopagnosic patients have suggested a double-dissociation between the processing of faces and other objects (Moscovitch, 1997). Also, neuroimaging studies have shown more activity for faces than objects from other categories in specific associative areas of the visual cortex (Kanwisher, McDermott, & Chun, 1997; Liu, 2008; Xu, Liu, & Kanwisher, 2005).

However, other authors have proposed an expertise-based account for N170 face-selectivity (Bukach, et al., 2006; Gauthier & Curby, 2005; Rossion, Curran, & Gauthier, 2002; Rossion, Gauthier, Goffaux, Tarr, & Crommelinck, 2002). The N170 component is particularly sensitive to faces due to the extensive experience that humans have with faces compared to other object categories. Participants trained to become experts with novel objects (e.g., greebles; Gauthier & Tarr, 1997) display greater N170 amplitude to these objects, even though it remains consistently larger for faces (Rossion, et al., 2000; Rossion, Kung, & Tarr, 2004). Furthermore, Tanaka & Curran (2001) demonstrated that early stages of object perception can be modified by experience and learning. For instance, they observed enhanced N170 amplitude to pictures of dogs when testing dog experts, which was associated with perceptual learning.

More recently, Thierry, Martin, Downing, & Pegna (2007a) noticed that in a number of ERP studies contrasting faces with other objects, faces were presented full front while contrasting objects were presented with varying degrees of perceptual variability that was not formally controlled. They argued that this variance in stimuli might have induced the differences observed in N170 amplitude between faces and other objects. In their study testing potential effects of stimulus variability on the N170 face-selectivity, Thierry et al. (2007a) manipulated inter-stimulus perceptual variance (low and high ISPV) and object category (faces and cars) in a two-by-two experimental design. ISPV refers to the variability in size (i.e., apparent distance), width/height ratio (i.e., distortion), eccentricity (i.e., location in space), and orientation, all of which carry important cues for perceptual analysis beyond the level of low-level physical differences. The N170 was significantly modulated by ISPV but, critically, not by stimulus category. In fact, the N170 amplitude was the same for cars and faces when ISVP was low. Interestingly, the absence of N170 amplitude variation between full front faces and full front cars has been reported before (Rossion, et al., 2000; Schweinberger, et al., 2004; Schweinberger, Kaufmann, Moratti, Keil, & Burton, 2007). Furthermore, Thierry et al. (2007a) found evidence for category-selectivity in the P1 range, i.e., 70 ms before the N170 time window.

The methods used and the interpretations by Thierry et al. (2007a) have been vehemently criticised (Bentin, et al., 2007; but see Thierry, Martin, Downing, & Pegna, 2007b). Critically, amongst the claims put forward to challenge Thierry et al.'s (2007a) conclusions, Rossion & Jacques (2008) referred to results of an unpublished study which shows significantly greater N170 amplitudes for full front faces than full front cars. Moreover, Rossion & Jacques (2008) argued that "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". Considering that (1) Rossion & Jacques (in preparation) controlled for ISPV and still revealed a N170 face-specific modulation and (2) stimulus categorisation is often chosen as a task (see for instance Rossion & Jacques, 2008), we replicated Thierry et al. (2007a) using a one-back task and a categorisation task. The categorisation task required an overt response to every stimulus whereas the one-back task only required a response when stimuli were immediately repeated. Furthermore, the categorisation task only required the processing of stimulus category to make the correct distinction, whereas in the one-back task, participants had to focus on configural information to perform the task correctly.

The goals of this study were to (1) replicate Thierry et al.'s (2007a) results and (2) investigate task effects on the N170 amplitude. Based on

Thierry et al. (2007a), we expected to find significant effects of ISPV and no effect of category on N170 amplitude and the reverse pattern in the P1 range. We also expected no interaction between category and task effects in either the P1 or the N170 range.

Methods

Participants

Seventeen participants (mean age 20.2 ±1.4 years, 11 right-handed, 14 females) volunteered from Bangor University's student participation scheme. Informed consent was obtained from all participants and the study was approved by the ethics committee at Bangor University. All participants had normal or corrected to normal vision.

Stimuli

Stimuli were 400 pictures divided in four equal groups: (a) full-front faces, controlled for size and position; (b) faces which varied in size, point of view and position; (c) full-front cars, controlled for size and position; (d) cars which varied in size, point of view and position (cf. Figure 1). Images in conditions (a) and (c) were scaled to a predefined template to produce images that were of similar size, height-width ratio and centred on the screen so as to minimize ISPV. Images in conditions (b) and (d) were cars and faces viewed from different angles, displaced randomly off-centre and reduced or magnified so as to maximise ISPV. All images were transposed onto a medium grey background (50%) and presented in greyscale. All stimuli subtended approximately 6 degrees of visual angle on the experimental display monitor.

The range of picture sizes was 4.96^0 ± 1.2 horizontally and 6.22^0 ± 1.3 vertically in condition b and 7.85^0 ± 0.9 horizontally and 3.77^0± 0.8 vertically in condition d The average luminance of all the images presented was 59.37cd/m^2 (SD = 27.71), calculated from average images for each condition (combined from all stimuli and sampled at 25 locations in a square grid for each image).

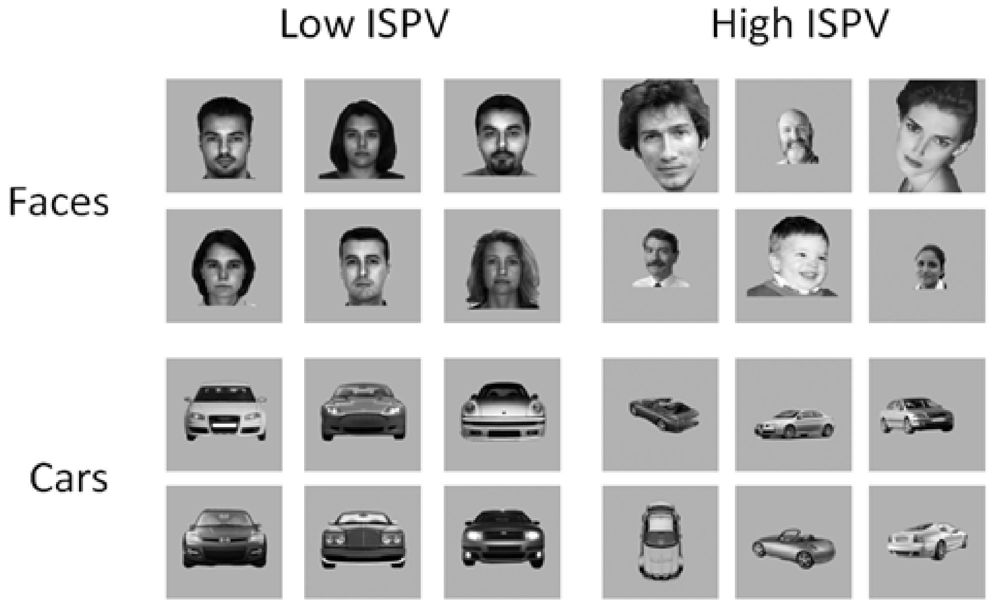

Figure 1 | **Experimental design. Examples of full-front faces and cars matched for size and eccentricity (low ISPV condition), and of faces and cars varying in size and eccentricity (high ISPV condition).**

Procedure

Participants were seated in a comfortable chair, 100 cm away from a low radiation TFT monitor. They were presented with a random stream of pictures from all four categories, each displayed for 200 ms with an inter-stimulus interval of 1500 ms in four blocks of 100 trials. Participants were asked to perform two tasks, a categorisation task and a one-back task, one after the other and on the same stimulus set. In the categorisation task, participants were asked to decide whether each image depicted a face or a car by pressing two keyboard keys set under their left and right index fingers. Response side was counter-balanced between participants. In the one-back task, eight pictures in each block were randomly selected and presented twice (repetition trial). Participants were then asked to press the space bar of a keyboard when they saw the exact same image twice in a row. Repeated images were removed from subsequent analyses. Task order was counterbalanced between participants.

ERP acquisition and processing

Electrophysiological data were recorded in reference to electrode Cz at a sampling rate of 1 KHz from 64 Ag/AgCl electrodes placed according to the extended 10-20 convention. Impedances were kept below 5 kOhms. The electroencephalogram was filtered on-line with a band pass filter between 0.01 Hz and 200 Hz, and off-line with a zero phase shift low pass digital filter set at 35Hz (48 db/octave slope). Eye blink artifacts were mathematically corrected based on a model artifact computed from a minimum of 50 individual artifacts in each participant using the procedure implemented in Scan 4.3 (Neuroscan, Inc., El Paso, TX, USA) and signals exceeding ±100 µV in any given epoch were automatically discarded. Continuous recordings were cut into epochs ranging from -100 to 400 ms after stimulus onset. Baseline correction was performed with reference to pre-stimulus activity and individual averages were digitally re-referenced to the global average reference. In the categorization task, epochs corresponding to errors were removed. In the one-back task, repetition trials were systematically excluded from averaging.

Statistical analysis

Participant's error rates and ERP measures were subjected to repeated-measures analyses of variance (ANOVAs). All comparisons were two-tailed, and the level of significance was set to 95% (p = 0.05). P1 and N170 mean amplitude analyses were conducted at electrode sites of maximal amplitude based on visual inspection for all conditions. The P1 was maximal at PO8, and mean amplitude analyses of P1 were conducted in a 20 ms time window around this peak (93 – 113 ms) at electrode sites O1, O2, PO7, PO8, PO9 & PO10. The N170 was maximal at PO10, and mean amplitude analyses were conducted in a 50 ms time window around the peak, between 130–180 ms. N170 peak amplitude was analyzed at 10 parietooccipital electrode sites (PO3, PO4, PO7, PO8, PO9, PO10, POZ, O1, O2, & OZ) to replicate Thierry et al. (2007a). A further analysis of N170 mean amplitude was conducted on selected electrode sites P7, P8, PO7, PO8, PO9 & PO10 considered "optimal" for measuring N170 face-selectivity (Rossion & Jacques, 2008). P1 peak amplitude and latency were subjected to a 2 x 2 x 2 x 2 x 6 repeated measures ANOVA. Factors were task (categorisation vs. one-back), category (face vs. car), ISPV (low vs. high), hemisphere (left vs. right) and electrode (6 levels). N170 peak amplitude and latency were subjected to a 2 x 2 x 2 x 10 repeated measures ANOVA in the first analysis. Factors were task (categorisation vs. one-back), category (face vs. car), ISPV (low vs. high) and electrode (10 levels). In the second analysis, N170 peak amplitude and latency were subjected to a 2 x 2 x 2 x 2 x 3 repeated measures ANOVA. Factors were task (categorisation vs. one-back), category (face vs. car), ISPV (low vs. high), hemisphere (left vs. right) and electrode (3 levels).

Results

Behavioural results. In the one-back task, mean accuracy was 100% and the rate of false alarms was 0%. Reaction times were not relevant in this task. In the discrimination task, mean accuracy was 94.0% ±5.6 and reaction times were 376 ±22 ms. Neither reaction times nor accuracy were affected by stimulus category or ISPV in the discrimination task (all Ps >.05).

ERP results. ERPs displayed a classic P1-N1-P2 complex in all experimental conditions (cf. Figure 2). Repeated measures analysis of variance (ANOVA) performed on P1 amplitudes over 6 parietooccipital electrodes revealed a main effect of object category (F[1,16] = 13.92, p = .002) such that P1 was significantly larger for faces than cars (Fig. 2). There was a significant task effect (F[1,16] = 9.80, p = .006) such that P1 was larger in the one-back than in the categorization task. There was a significant hemisphere effect (F[1,16] = 5.84, p = .03) such that P1 was significantly larger in the right than left hemisphere. There was no significant main effect of ISPV (F[1,16] = 1.48, p = .24). None of the interactions were significant (all Ps > .1). None of the experimental factors elicited significant effects on P1 latencies (Category: F[1,16] = 2.10, p = .17; ISPV: F[1,16] = .23, p = .64; Task: F[1,16] = .001, p = .98) and no interaction was significant (all Ps > .10).

A repeated measures ANOVA performed on N170 amplitudes over 10 electrodes revealed a significant effect of ISPV (F[1,16] = 11.56, p = .004) such that N170 amplitude was larger in the low ISPV condition. There was no effect of object category (F[1,16] = .60, p = .45) and no task effect (F[1,16] < .001, p = .99). No interaction was significant (all Ps > .25). The latency of the N170 was delayed for cars as compared to faces (F[1,16] = 74.29, p < .0001). There was a significant ISPV effect (F[1,16] = 36.06, p < .0001) and a significant interaction between ISPV and category (F[1,16] = 20.49, p < .001), such that N170 peak latency was increased more greatly for high as compared to low ISPV for cars than for faces. There was no effect of task (F[1,16] = 1.51, p = .24) on the N170 latency and no other interaction was significant (all Ps > .3).

The second analysis of the N170 mean amplitude performed on electrodes P7, P8, PO7, PO8, PO9 & PO10 (2 x 2 x 2 x2 x 3 analysis) revealed a main effect of category (F[1,16] = 5.01, p = .04) such that the N170 was significantly larger for cars than faces (cf. Figure 2). ISPV significantly affected the N170 mean amplitude (F[1,16] = 31.89, p < .0001) and interacted with category such that the effect of ISPV was greater for cars than faces (F[1,16] = 32.22, p < .0001). There was no main effect of task (F[1,16] = .04, p = .84) nor hemisphere (F[1,16] = .96, p = .34). There was a marginally significant hemisphere x category interaction (F[1,16] = 4.52, p = .05) showing that the category effect was larger over the right hemisphere. No other interaction was significant (all Ps > .1).

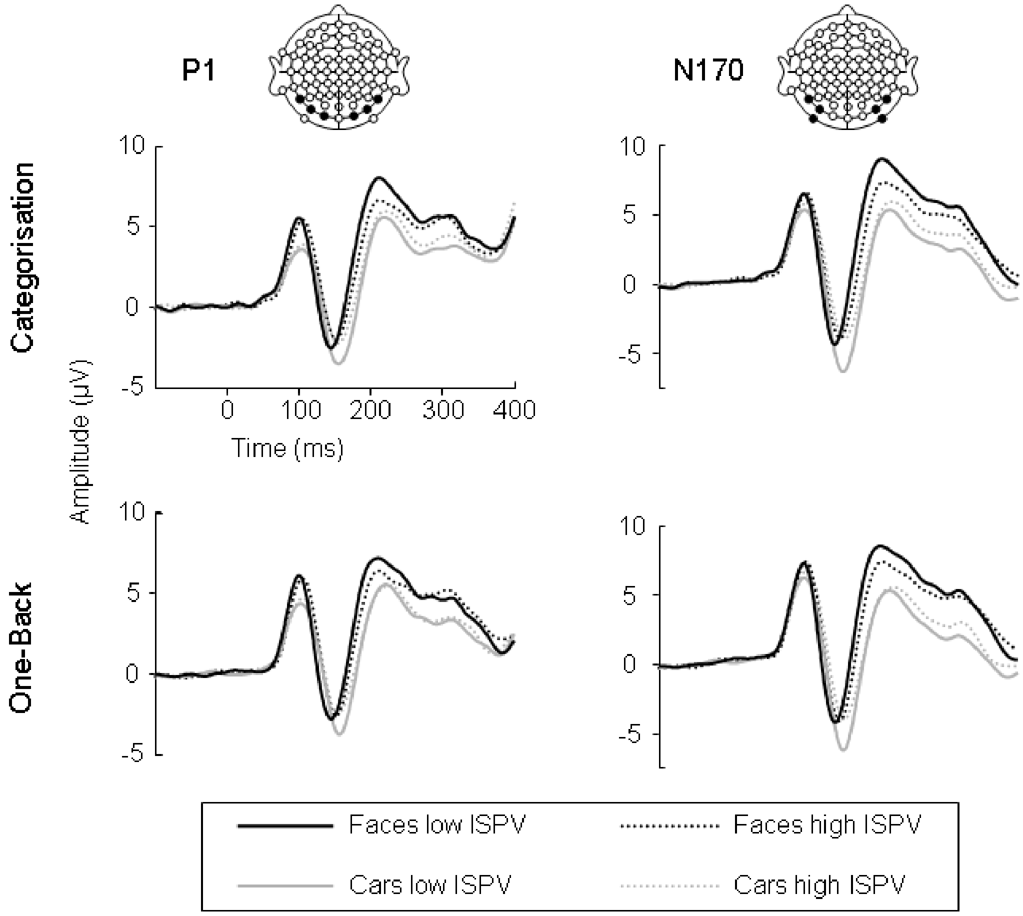

Figure 2 | **Grand average event-related potentials recorded in the four experimental conditions in each of the two tasks. Waveforms depict a linear derivation of the electrodes used in statistical analyses for the P1 and N170, respectively.**

Discussion

P1 results replicated Thierry et al. (2007a): P1 was modulated by object category and not by ISPV (Fig. 2). Although these results need to be replicated with control object categories other than cars, they suggest again that categorical effects might be observed earlier than the N170 range.

As expected and consistent with previous findings (Rossion, et al., 2000; Schweinberger, et al., 2004; Schweinberger, et al., 2007; Thierry, et al., 2007a), no significant amplitude differences were found in the N170 range between full front views of faces and cars across ten parietooccipital electrode sites. This result replicates Thierry et al. (2007a). When the N170 was analyzed at electrode sites suggested by Rossion & Jacques (2008), a significant category effect appeared, however, but opposite to the "established" effect, i.e., N170 amplitude was significantly greater for cars than faces. This result stands in contrast to the majority of publications reporting N170 category-selective effects (Bentin, et al., 1996; Carmel & Bentin, 2002; Rossion, Curran, et al., 2002; Rossion, Gauthier, et al., 2002; Rossion, et al., 2004), in which interstimulus perceptual variance (ISPV) was not formally controlled.

Furthermore, since the category effect was significant, this result cannot simply be dismissed as a non-result as argued by Bentin et al. (2007). This result further challenges the conceptualisation of the N170 as a face-selective marker. Note that this observation only concerns N170 amplitude since the latency results consistently show a delay for cars as compared to faces (Rossion, et al., 2000; Thierry, et al., 2007a).

As in Thierry et al. (2007a), ISPV had a significant effect on N170 mean amplitude, i.e., low ISPV stimuli elicited greater N170 activity than high ISPV stimuli. The interaction between category and ISPV was reported previously (Thierry, et al., 2007a) and is probably due to greater variability in the case of high ISPV cars, which were shown from a variety of different viewpoints, whereas faces always displayed most internal features (there were no profile or rear views of faces).

Interestingly, no main task effect was observed on the N170 amplitude and latency. The N170 event did not significantly differ between the one-back and categorisation task contexts, showing that Thierry et al.'s (2007a) results cannot be accounted for by a task effect. There was however a main effect of task on P1 amplitude, indicating that the one-back task may have involved higher attentional load (Taylor, 2002) or higher involvement of selective attention (Hopf & Mangun, 2000).

In their critical review of Thierry et al. (2007a), Rossion & Jacques (2008) introduced new data from a study in which they compared cars and faces from a full-front viewpoint. They reported significantly greater amplitudes to faces than cars, a result which stands in direct opposition with those of the present study. Unfortunately, it is impossible to consider the full range of factors that may have accounted for the difference between the two studies because data presented by the authors are not yet published and the full methodological details of this study are therefore unavailable. From the information the authors provide in their critique, we can only speculate that some parameters such as stimulus repetition and specific properties of the stimuli used (e.g., low-level contrast, absolute size, number of repetition of stimuli from each category, use of detoured faces instead of full heads, etc.) account for the striking difference in experimental outcome. Further investigations will be needed to determine which of the experimental parameters listed above affect N170 amplitude and why.

Conclusion

We found that full-front views of cars elicit significantly larger N170 than faces at electrode sites where face-selectivity is classically considered optimal (Rossion & Jacques, 2008). These results further challenge the face-selectivity of the N170. Moreover, the absence of face-selectivity in the N170 range is unlikely to merely result from the task used since we found no difference in a one-back and a categorisation task.

Chapter 3

Category-sensitivity in the N170 range: A question of topography and inversion, not one of amplitude

This chapter is published as: Boehm, S. G.; Dering, B.; & Thierry, G. (2011). Category-sensitivity in the N170 range: A question of topography and inversion, not one of amplitude. *Neuropsychologia,* 49(7), 2082 – 2089.

Abstract

Event-related potential studies have identified the N170 as the key neurophysiological marker of human face processing. This functional association relies on the observation of a larger N170 amplitude to faces than items from all other visual object categories. However, N170 amplitude is modulated by stimulus variations like viewpoint, size and symmetry, and studies comparing similarly sized and symmetric full-front faces and other objects have failed to find amplitude differences. Here we tested whether the effect of inversion – an increase in N170 amplitude seen for faces presented upside down – is similarly observed for full-front views of cars. Participants discriminated pictures of faces and cars, which were presented upright and inverted, and either in full-front view or varying in size, orientation and viewpoint. For upright stimuli, the N170 was stronger for faces than cars at some electrode sites, but of comparable amplitude at others, as shown by topographical differences. The N170 for inverted faces and cars was delayed, with a stronger delay for faces than cars. Inversion increased N170 amplitude for faces, while modulations for full-front view cars were non-significant or N170 amplitude was reduced. These results further limit the widely acknowledged principle of an association between N170 and visual object categorization. Potential face-sensitivity in the N170 range may therefore rely on topographic differences and effects of inversion, rather than amplitude differences.

Introduction

Over the past fifteen years, the N170 has become the most widely used neurophysiological marker of face processing. This peak of event-related potentials (ERPs) is characterized by bilateral temporal negative deflections at around 170ms after stimulus onset and a corresponding positive deflection maximal at the vertex (Bentin, et al., 1996; Eimer, 2000b; Jeffreys, 1996). The N170 and the M170, its magnetic counterpart (Liu, Higuchi, Marantz, & Kanwisher, 2000; Xu, et al., 2005), are thought to index structural encoding of faces, like configural processing (Eimer, 2000b; Mercure, Dick, & Johnson, 2008; Sagiv & Bentin, 2001). Thus the N170 has been used, for example, to study the neural bases of prosopagnosia (Bentin, Deouell, & Soroker, 1999; Eimer & McCarthy, 1999; Minnebusch, Suchan, Ramon, & Daum, 2007). The N170 has been shown to be generally increased in amplitude for faces when compared to many other categories of objects (Eimer, 2000b; Itier, Latinus, & Taylor, 2006; Itier & Taylor, 2004b).

From this larger amplitude to faces, it is commonly concluded that the N170 is face-sensitive, that is, it is thought to reflect processes absent in the case of other visual stimulus categories (Rossion & Jacques, 2008). Faces and other stimulus categories vary based on a number of visual properties such as luminance, contrast, frequency, orientation, symmetry, and size, some of which may modulate N170 amplitude. Thierry, Martin, Downing, and Pegna (2007) manipulated some of these properties, comparing the processing of faces, cars and butterflies, either similar or variable in terms of size, orientation, and symmetry across experimental trials. When faces and other categories were presented in full-front views with similar size, the larger N170/M170 amplitude to faces vanished (Schweinberger, Huddy, & Burton, 2004; Schweinberger, Kaufmann, Moratti, Keil, & Burton, 2007; Thierry et al., 2007a).

The P1 is also sensitive to some of the same visual stimulus properties as the N170, and can be similarly influenced by the same variations (Dering, Martin, & Thierry, 2009). In contrast to the N170, P1 has been significantly modulated by category such that P1 amplitude was larger for faces compared to other categories (Dering et al., 2009; Thierry et al., 2007a). The association of P1 with category-sensitivity for faces (Dering et al., 2009; Herrmann, Ehlis, Ellgring, & Fallgatter, 2005; Liu, Harris, & Kanwisher, 2002; Thierry et al., 2007a), however, is far from being commonly acknowledged (Rossion & Jacques, 2008). Category differences sometimes extend beyond the N170, for example encompassing the P2, and face-sensitive processing has been shown in later time ranges (Boehm, Klostermann, & Paller, 2006; Milivojevic, Clapp, Johnson, & Corballis, 2003; Nasr & Esteky, 2009).

One of the most prominent effects associated with face processing is the so-called inversion effect. It is well established that face inversion impairs face recognition (Yin, 1969) by impoverishing configural processing of first and second order relations, as well as holistic processing (Maurer, Le Grand, & Mondloch, 2002). Interestingly, the N170 is commonly affected by face inversion: its amplitude is increased and its latency is delayed when faces are presented upside-down (Bentin et al., 1996; Boehm et al., 2006; Eimer, 2000a; Itier et al., 2006; Itier & Taylor, 2002, 2004a; Jeffreys, 1996). A similar increased amplitude of the N170 can be observed for houses and words (Eimer, 2000a; Itier et al., 2006; Rossion, Joyce, Cottrell, & Tarr, 2003), while for ape faces, eyes in isolation, and greebles, inversion induces opposite or null effects on N170 amplitude (Itier, et al., 2006; Rossion, et al., 2000; Rossion, Joyce, Cottrell, & Tarr, 2003; Weise, 2009). Despite some inconsistencies in the literature, the consensus is that the inversion effect on the N170 is similarly unique to faces as the behavioural inversion effect (Bentin et al., 2007).

The goal of the present study was to investigate whether the effect of stimulus variability (suppressing apparent face-sensitivity when stimuli are presented in full-front views) on the N170 to upright stimuli, extends to the inversion effect. In other words, will we find similarly increased N170 amplitude by inversion for full-front views of cars? If inversion affects uniquely configural processing, the effect of inversion on N170 should remain face-sensitive even when faces and cars are presented with similar sizes and in full-front views. On the other hand, an inversion effect comparable for faces and cars would further question the face-sensitivity of the N170.

We presented faces and cars in a two (face/car) by two (full-front/variable views) by two (upright/inverted) experimental design. Participants indicated the category of each picture presented by button presses. We analyzed P1, P2 and N170 with a focus on the influence of stimulus variability (full-front vs. views variable in size, viewpoint and symmetry), inversion, and the interaction between stimulus variability and inversion.

Methods

Participants

Thirty-four young adults participated for course and print credit. The data from five participants were discarded because a clear P1 or N170 was not detectable in one or more conditions. The mean age of the remaining participants (18 females) was 22 years (range 18–48). All participants had normal or corrected-to-normal vision, and 25 were right-handed as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971). The study was approved by the ethics committee at the School of Psychology, Bangor University, and all participants gave written informed consent to take part in the study.

Stimuli

Upright stimuli were taken from Thierry et al. (2007a) and consisted of one hundred greyscale pictures for each of the following groups: (a) full-front view faces; (b) faces varying in size, orientation, position, and individual's age, (c) full-front view cars, (d) cars varying in size, orientation, position, make and manufacture period (Fig. 1). Faces and cars were extracted and transposed on a medium grey background. In the case of groups (a) and (c), images fitted a predefined template of size, viewpoint, and had similar levels of symmetry. In the case of groups (b) and (d), half of the pictures were larger and half were smaller than the pictures in (a) and (c), balancing average stimulus size. Full-front view faces and cars were used for the full-front view conditions, and variable cars and faces for the variable conditions. In addition to the original stimuli from Thierry et al., all pictures were also presented upside-down (inverted conditions).

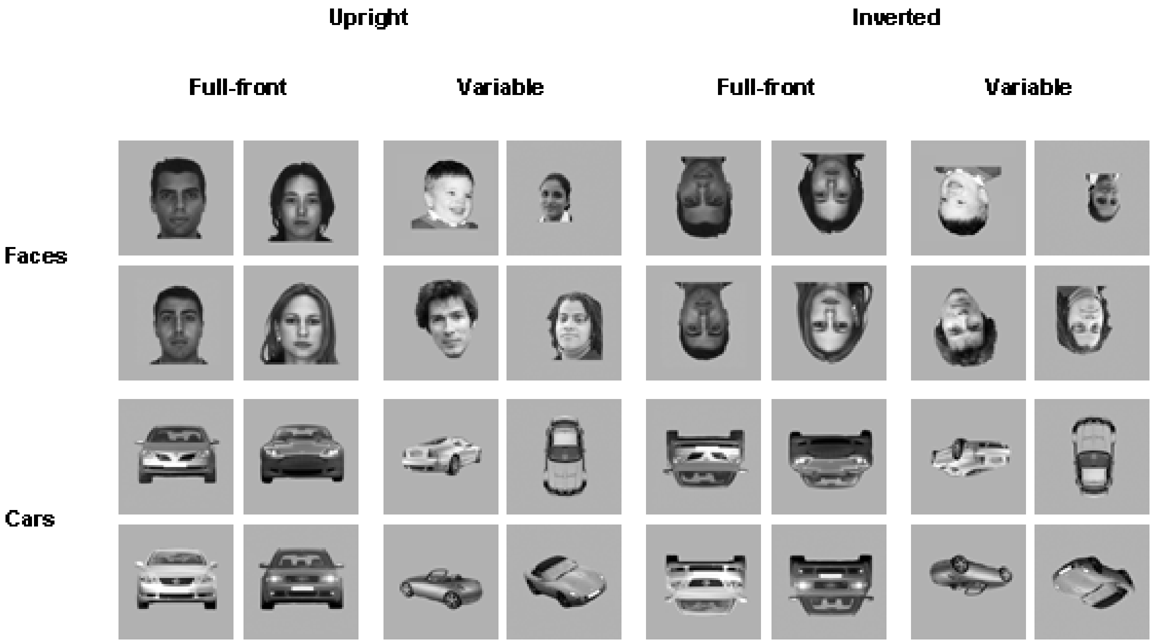

Figure 1 **Stimulus exemplars from the experimental conditions. Inverted conditions comprise the same stimuli presented upside down.**

Procedure

Participants sat in a comfortable chair in front of an LCD computer screen in a dimly lit room. Stimuli were presented for 200 ms separated by a fixation cross for 1500 ms in pseudo-randomized order in 8 blocks of 100, such that each block featured 12 or 13 pictures from each of the eight experimental conditions. It was therefore impossible for participants to predict what condition would be presented from one trial to the next. All stimuli subtended approximately 6^0 of visual angle on the experimental display monitor. The range of sizes in variable conditions were 4.96^0 ±1.2 horizontally and 6.22^0 ±1.3 vertically for faces and 7.85^0 ±0.9 horizontally and 3.77^0 ±0.8 vertically for cars. Participants were asked to judge the category (face, car) of each stimulus presented and press the “F” and “J” key with the index fingers of their left and right hand for faces or cars, irrespective of orientation. The assignment of keys/hands to responses (faces/cars) was counterbalanced across participants. Instructions emphasized both speed and accuracy. Correctness of response and response time were measured; missing responses were treated as errors.

EEG recording, ERP methods

The continuous electroencephalogram (EEG) was recorded from 64 scalp positions according to an extended 10–20 system from Ag/AgCl electrodes mounted in an elastic cap, referenced against Cz as initial common reference and digitized with a frequency of 1 kHz. The band-pass was set to 0.01–200Hz. Horizontal and vertical electrooculograms (EOG) were also recorded. Electrode impedances were kept below 5 kΩ.

The EEG was filtered offline with a low pass filter set at 35Hz and a slope of 48 dB/octave (zero phase shift). Eye blink artifacts were corrected mathematically (Gratton, Coles, & Donchin, 1983), and the EEG was separated into epochs of 800 ms length, starting 200 ms before stimulus onset. All epochs with correct responses were averaged into ERPs for each of the eight experimental conditions and each channel. The 200 ms pre-stimulus interval served as a baseline. The ERPs were rereferenced to the average reference in order to remove possible effects resulting from the choice of the reference electrode.

P1, P2 and N170 were quantified by means of peak amplitude and N170 also for peak latency, with peak latency measured in relation to stimulus onset and peak amplitude measured relative to baseline. Peaks for P1 were identified within 80–110 ms after stimulus onset at PO7 and PO8, for P2 within 180 and 300 ms at O1 and O2, and for N170 within 120–230 ms at PO9 and PO10. Electrode selection was guided by the largest amplitude of P1, P2, and N170 in the grand average of upright full-front view faces over all participants. Additionally, N170 peak amplitude was also analyzed at the often used sites P7 and P8 (data from two participants needed to be discarded for this analysis because no clear N170 was detectable in at least one condition), and at PO7, PO8, where the amplitude was strongest for upright full-front view cars. In addition, the vertex positivity, considered a mirror-image of the N170 (Joyce & Rossion, 2005), was analyzed at Cz. For all of these additional analyses, the same time window as for the N170 analyses at PO9 and PO10 was used (120–230 ms). All peaks were detected automatically as the most positive (P1, P2, vertex positivity) or most negative (N170) local extrema within the corresponding time interval. Note that therefore the local maxima (and minima) do not correspond necessarily to the most positive (or negative) point of the waveform in the given interval. Peaks were verified by visual inspection. Prior to peak detection, the relevant ERP waveforms were filtered with a low-pass filter at 20Hz and a slope of 48 dB/octave to reduce the influence of high-frequency noise (Picton, et al., 2000).

Localization of brain generators for the N170 of full-front view faces and cars was done with Brain Electrical Source Analysis software (BESA version 5.1; Megis, Munich). For a given EEG signal measured at the scalp, BESA estimates the underlying generators as current source dipoles in a standard brain. Note that for the inverse problem of finding a generator solution for a given EEG signal, an infinite number of solutions are possible. To achieve a unique solution, specific assumptions have to be employed. In the present case we assumed two dipoles for faces, one in each hemisphere, and two similar dipoles for cars, using a 4-shell ellipsoidal head model. In order to reduce the influence of residual noise, the localization was done on the basis of the average signal from all participants (Picton et al., 2000).

Statistical analysis

Mean reaction times, error rates, and ERP measures were subjected to repeated-measure analyses of variance (ANOVA). Topographies were compared by interactions between conditions and electrode site in an ANOVA (McCarthy &Wood, 1985). All comparisons were two-tailed. The level of significance was set to $\alpha= 0.05$. Huynh–Feldt corrections were calculated whenever appropriate; the uncorrected degrees of freedom, the corrected p-value and the calculated Huynh–Feldt ε are reported.

Results

Accuracy was high in the upright conditions (see Fig. 2). Inversion reduced accuracy for faces, but increased accuracy for cars, as shown in an interaction between category (faces, cars) and inversion (upright, inverted), F(1,28) = 9.34, p = .0049. Additionally, inversion prolonged responses to faces and speeded up responses to cars, F(1,28) = 9.36, p = .0048 (Fig. 2). The effect of inversion on accuracy and response times was not modulated by variability as indicated by the absence of two-way interactions between category, inversion and variability (full-front view, variable view), Fs(1,28) < = 0.89, ps > = .353. In sum, inversion reduced performance for faces, but increased performance for cars.

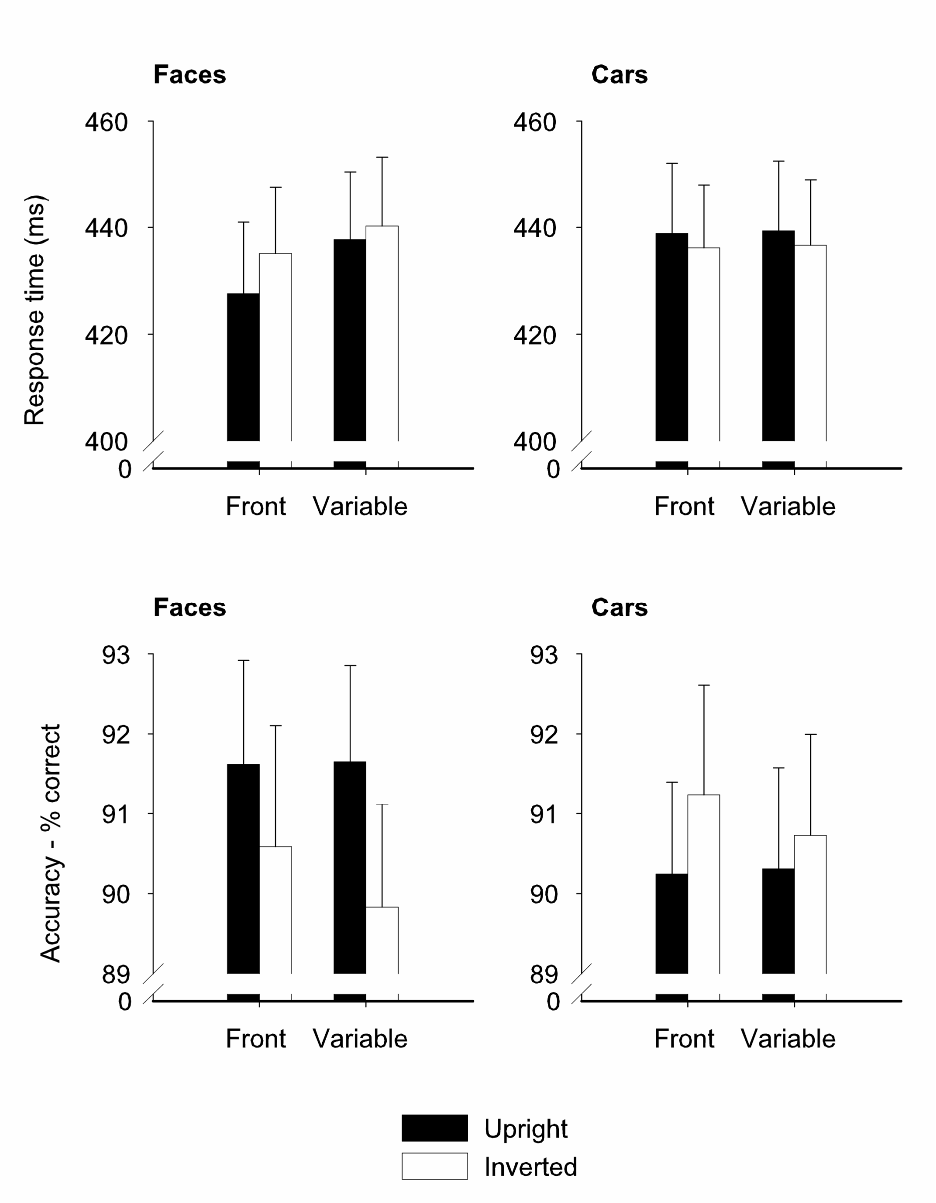

Figure 2 **Behavioural results. Accuracy and response times for the categorical judgment faces/cars, in relation to category, stimulus variability, and stimulus orientation. Error bars depict the standard error of the mean.**

Electrophysiological results

ERPs displayed a characteristic pattern with a P1 peaking around 105 ms at occipital electrodes, an N170 peaking around 150 ms at lateral posterior temporal electrode sites and a vertex positivity at central sites, followed by a P2 peaking around 240 ms. Specifically, clear N170 and vertex positivity were present for both faces and cars in all upright conditions (Fig. 3). The latency of the N170 seemed shorter for faces than cars. Inversion delayed the N170 for both faces and cars, but inversion had opposite effects across categories on N170 amplitudes (Fig. 4): N170 was more negative for inverted than upright faces, but less negative for inverted than upright cars. The ERP results were further quantified by ANOVA with a focus on the questions of category sensitivity (faces vs. cars) and stimulus variability (full-front vs. variable view) in P1, P2,N170 and vertex positivity, and on inversion (upright vs. upside down) in the N170 and vertex positivity.

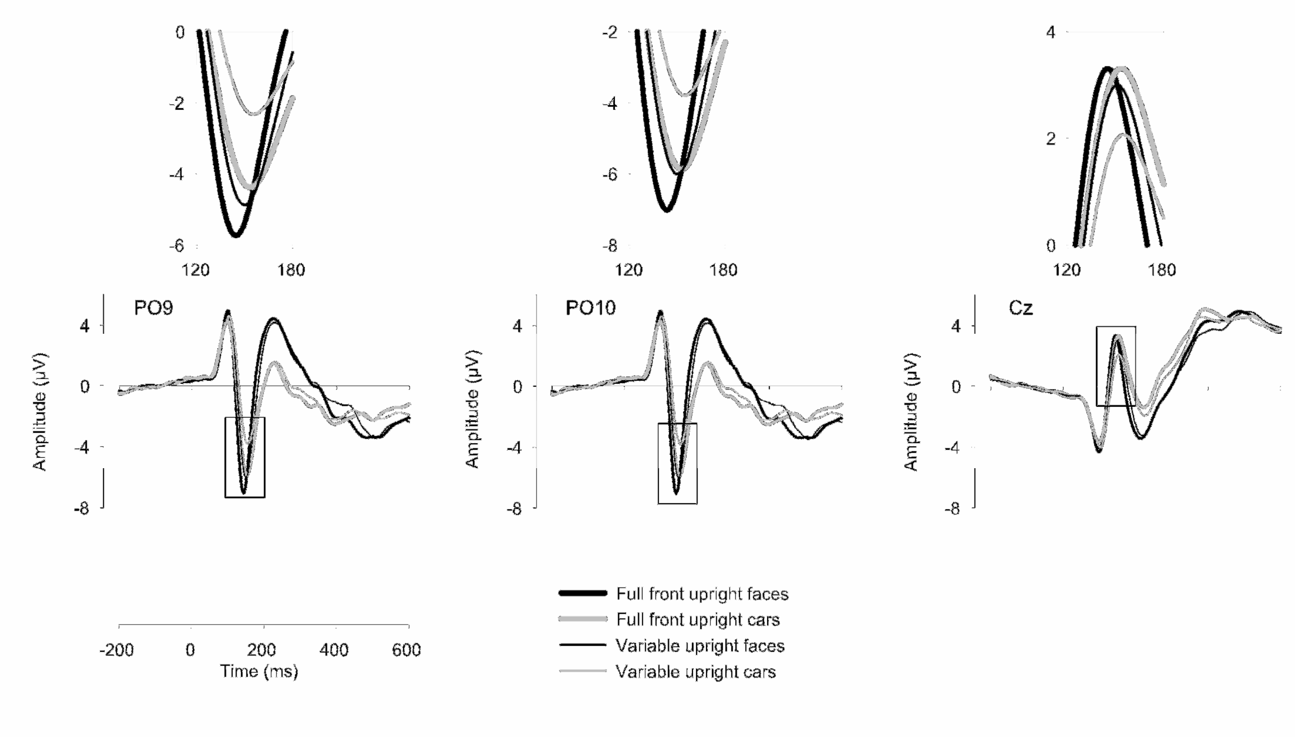

Figure 3 **ERP waveforms for upright faces and cars at selected electrode sites for N170 and vertex positivity.**

Upright stimuli

P1 peak amplitude was larger for faces than cars, F(1,28) = 19.33, p = .0001, and smaller for full-front than variable view stimuli, F(1,28) = 5.71, p = .0239. The category difference was not modulated by stimulus variability or vice versa, F(1,28) = 0.08, p = .7809. These results indicate that P1 is larger for faces and variable view stimuli.

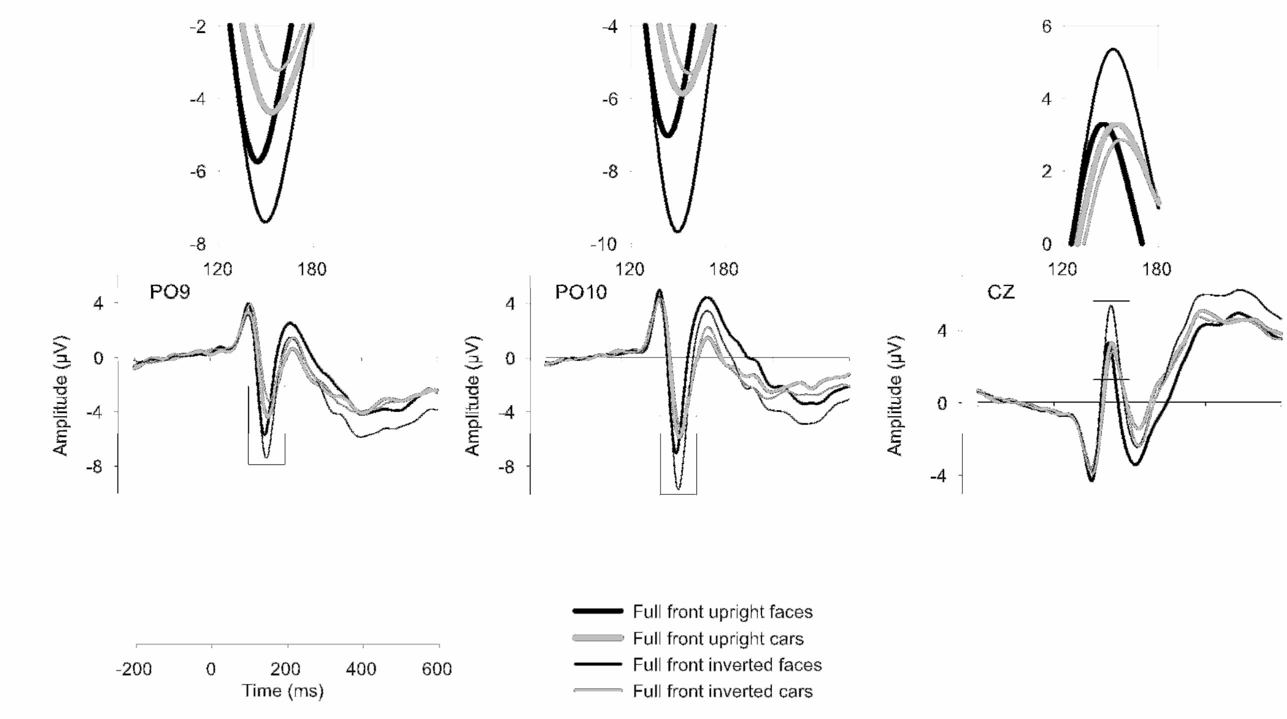

Figure 4. **ERP waveforms of upright and inverted faces and cars (full-front views only) at selected electrode sites for N170 and vertex positivity**

The P2 was larger for faces than cars, F(1,28) = 12.08, p = .0017, and smaller for full-front than variable view stimuli, F(1,28) = 5.41, p = .0274. The category difference was not modulated by stimulus variability or vice versa, F(1,28) = 2.25, p = .1450. These results indicate a larger P2 for faces and variable view stimuli.

At PO9 and PO10, where the N170 was strongest for faces, N170 amplitude was larger (i.e., more negative) for faces than cars, F(1,28) = 18.44, p = .0002 (Fig. 5). N170 amplitude was also more negative for full-front view than variable view stimuli, F(1,28) = 66.19, p < .0001. In addition, the amplitude difference between full-front and variable view stimuli was smaller for faces than cars, as indicated by a significant interaction of category and stimulus variability, F(1,28) = 8.16, p = .0080. The contrast usually reported in the literature between full-front view faces and variable cars showed a larger, that is more negativeN170 amplitude for faces, F(1,28) = 62.19, p < .0001. The N170 was still more negative for full-front views of faces and cars, F(1,28) = 7.73, p = .0096. These results indicate a larger, more negative N170 for full-front versus variable view stimuli, a weaker influence of stimulus variability for faces than cars, and a larger N170 for faces than cars.

When the N170 was analyzed at the often used sites P7 and P8, several of the findings found at PO9 and PO10 were replicated (Fig. 5). Again, N170 amplitude was more negative for full-front than variable view stimuli, F(1,26) = 46.81, p < .0001, and the difference between full-front and variable view stimuli was smaller for faces than cars, as indicated by a significant interaction of category and stimulus variability, F(1,26) = 27.62, p < .0001. Also, the comparison between full-front view faces and variable view cars showed a larger, more negative N170 amplitude for faces than cars, F(1,26) = 8.41, p = .0075. In contrast to the findings at P09 and PO10, however, N170 amplitude was comparable for faces and cars, F(1,26) = 0.00, p = .9901.

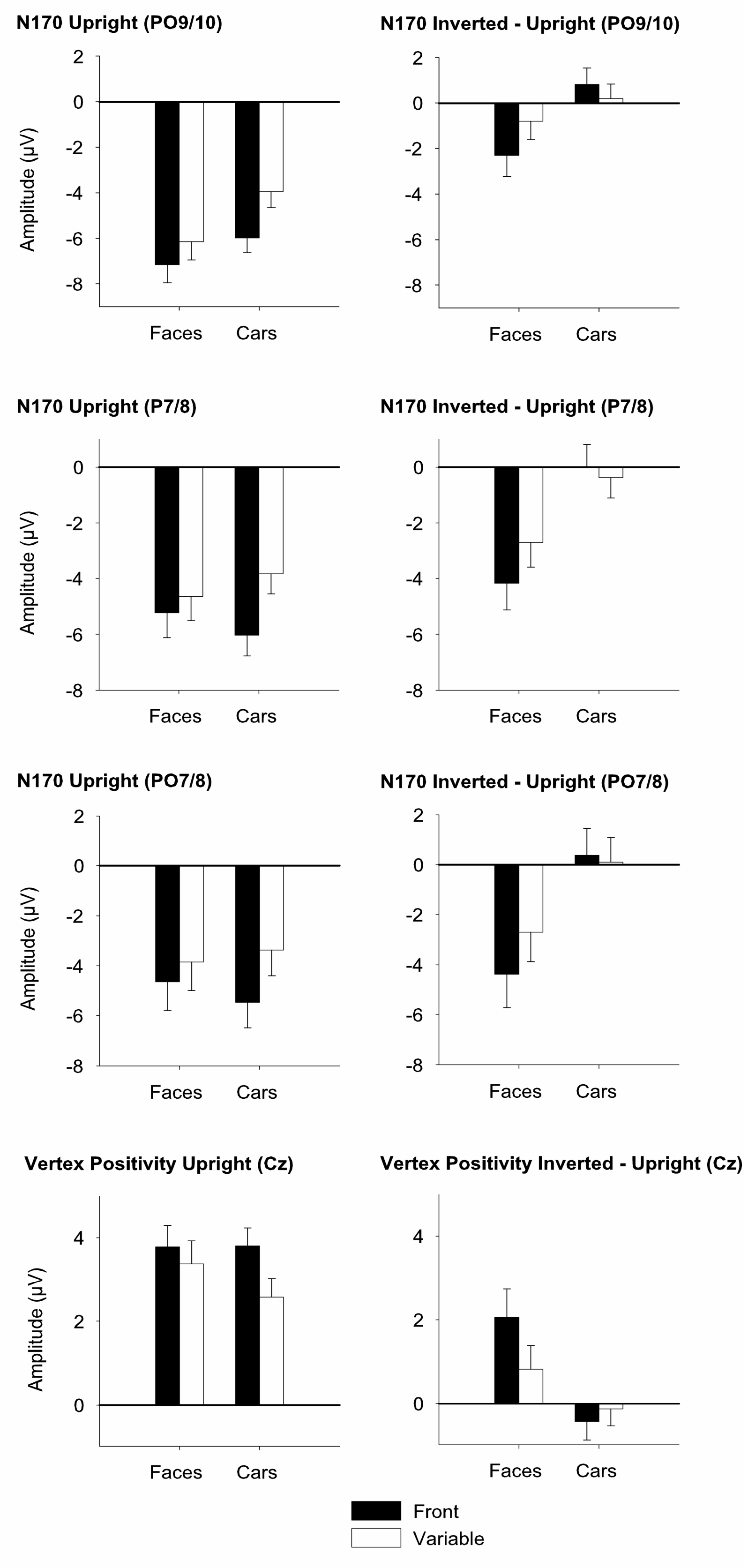

Figure 5 **Mean peak amplitudes for upright stimuli and mean peak amplitude differences between inverted and upright stimuli for N170 and vertex positivity at**

selected electrode sites in relation to category, stimulus variability and inversion. Error bars depict the standard error of the mean.

In the full-front view conditions, the N170 appeared to be smaller to faces than cars, but this difference did not reach significance, F(1,26) = 2.89, p = .1012. Thus the larger N170 for faces found at PO9/PO10 did not extend to P7 and P8.

The partially different findings in relation to the selection of electrode sites suggest topographic differences between the N170 elicited by faces and that elicited by cars (Fig. 6). Fortunately, the topographies for N170 for upright faces and upright cars in the full-front view conditions were of similar magnitude as measured by the mean amplitude across all electrodes except EOG and average reference sites, so no scaling (McCarthy & Wood, 1985) was required. The topographies were found indeed to differ in distribution, F(62, 1736) = 2.50, p = .0424, ε = .0687, indicating qualitative processing differences between theN170 to faces and cars. A source localisation with BESA revealed that the dipoles for faces were located more inferior than the dipoles for cars and that the dipoles in the left hemisphere had opposite orientations (Fig. 6).

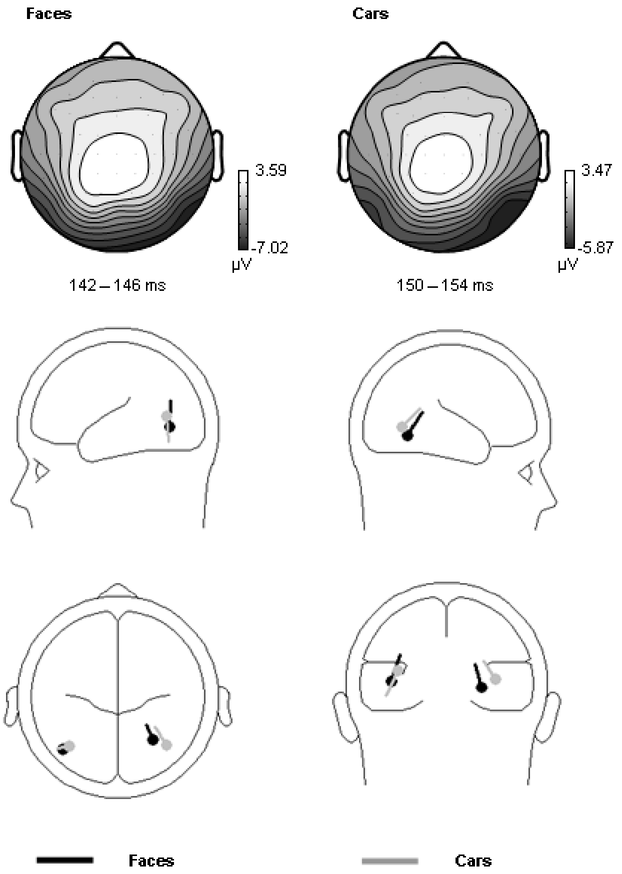

Figure 6. **Topographic shape maps of the N170 for upright full-front view faces and cars and the corresponding locations of the current source dipoles.**

Given this topographic difference between the N170 to faces and cars, N170 was also analyzed at PO7 and PO8, where it was strongest for full-front view cars (Fig. 5). N170 amplitude was comparable for faces and cars, F(1,28) = 0.14, $p = .7117$, more negative for full-front versus variable view stimuli, F(1,28) = 37.26, $p < .0001$,and the difference between full-front versus variable view stimuli was smaller for faces than cars, F(1,28) = 14.07, $p = .0008$. N170 amplitude was larger for full-front view faces than variable view cars, F(1,28) = 5.11, $p < .0317$. Importantly, although in the full-front view conditions N170 appeared to be smaller for faces than cars, this difference did not reach significance, F(1,28) = 2.81, p = .1046. Hence, the results obtained at PO7 and PO8 mirrored those found at P7 and P8; that is, they replicated some of the findings obtained at PO9 and PO10, but contrasted with the larger N170 for faces at PO9 and PO10.

The amplitude of the vertex positivity (Fig. 5) was comparable for faces and cars, F(1,28) = 2.95, p = .0971, more positive for full-front than variable view stimuli, F(1,28) = 36.37, p < .0001, and the difference between full-front and variable view stimuli was smaller for faces than cars, F(1,28) = 8.14, p = .0080. Amplitude was larger for full-front view faces than variable view cars, F(1,28) = 20.56, p < .0001, but did not differ between full-front view faces and full-front view cars, F(1,28) = 0.01, p = .9258. Thus, the vertex positivity mirrored the results found at P7, P8, PO7 and PO8, but did not show the larger amplitude to faces found at PO9 and PO10.

The latency of the N170, measured at PO9 and PO10,was generally shorter for faces than cars, and shorter for full-front view versus variable view stimuli, Fs(1,28) > = 37.00, ps < .0001.Additionally, the latency shortening by full-front view versus variable view stimuli was larger for faces than cars, F(1,28) = 15.90, p = .0004.

Inversion effects

The effect of stimulus inversion on the peak amplitudes of N170 and vertex positivity was analyzed as the difference in peak amplitude between the inverted and the corresponding upright conditions. For the N170, a negative difference indicates a more negative, that is, a more pronounced N170 for inverted than upright stimuli, whereas for the vertex positivity, a positive difference indicates a more pronounced vertex positivity for inverted than upright stimuli (Fig. 4). Across all selected electrodes (PO9/10, P7/8, PO7/8, Cz), inversion resulted in an enhanced N170 and vertex positivity for faces, whereas for cars no consistent enhancement of N170 and vertex positivity was found, as indicated by interactions between category and inversion, Fs(1,26/28) > = 39.00, ps < .0001 (Fig. 5). More specifically, the effect of inversion for cars was negligible at most sites, Fs(1,26/28) < = 1.73, ps > = .1995, except for a reduction at PO9/10 and Cz, Fs(1,28) > = 4.44, ps < = .0441. Full-front view compared to variable view stimuli showed stronger effects of inversion; this was shown in a significant interaction of category and stimulus variability, Fs(1,26/28) > = 9.07, ps < = .0055, because inversion induced changes for faces and cars in opposite directions. The overall pattern of enhanced N170 for faces and negligibly modulated or reduced N170 for cars was also present in the usual comparison between full-front view faces and variable view cars, Fs(1,26/28) > = 31.21, ps < .0001, and this effect was not abolished for full-front view stimuli in both categories, Fs(1,26/28) > = 33.64, ps < .0001. In sum, these findings indicate that the inversion effect is category-sensitive, the N170 and the vertex positivity for inverted in comparison to upright faces is enhanced, but not modulated or reduced for inverted in comparison to upright cars. Moreover, the face-sensitivity of the inversion effect onN170 amplitude remained present for full-front view stimuli.

The effect of stimulus inversion on N170 peak latency was analyzed similarly to the analyses of inversion on N170 amplitude as the difference in peak amplitude between the inverted and the corresponding upright conditions. The peak of the N170 was delayed for both full-front and variable view faces, Fs(1,28) > = 93.58, ps < .0001. Although inversion delayed the peak of the N170 stronger for faces than cars, F(1,28) = 8.46, p = .0070, clear delays by inversion were present for both full-front and variable view cars, Fs(1,28) > = 16.36, ps < = .0004. N170 latency was not influenced by stimulus variability, nor did stimulus variability modulate the peak delay by inversion, Fs(1,28) < = 1.57, ps > = .2211. Inversion affected N170 latency for both categories, but stronger for faces than cars.

Discussion

This study aimed at testing the effect of stimulus inversion on ERP markers of visual processing under varying conditions of size, orientation and symmetry with a strong focus on the N170. At the most inferior temporal electrode sites PO9 and PO10, where the N170 was maximal in amplitude for faces, the N170 was clearly present for cars, although somewhat smaller than for faces. The N170, in contrast, was similar for cars and faces at the commonly used sites P7 and P8, as well as at PO7 and PO8, where the N170 was maximal for cars. Inversion delayed the N170 peak for faces and cars, with the delay being larger in the case of faces than cars. The increased amplitude of the N170 by inversion was found only for faces, whereas car inversion negligibly modulated or reduced N170 amplitude. Finally, the topography of the N170 for upright full-front view stimuli was different for faces and cars.

The present results replicate P1 sensitivity to category shown by Thierry et al. (200a7). The insensitivity of N170 to category (Thierry et al., 2007a) was replicated at some electrodes, but not others. These prior results have been used to argue against the common view that the N170, but not the P1 is associated with category, and this critique is partly upheld by the present findings. In line with previous research (Schweinberger et al., 2004, 2007), the N170 for cars and faces was similar at both the sites where the amplitude for cars was maximal, as well as at the commonly used sites P7/8.

In contrast, the N170 was larger to faces than cars at PO9/10. At these electrodes, the N170 also showed its largest amplitude to faces. Thus, according to the present data, the N170 appears face sensitive when selecting electrode sites where it is maximal for faces. This face sensitivity does not extend to other sites, however. In addition, the N170 can be just as car sensitive with an equal bias of electrode selection towards cars (Dering et al., 2009). Based on the present results, the larger N170 to faces than other categories, commonly held in favour of the N170 to be face sensitive, may be – at least in part – a result of a selection bias of electrodes towards faces. Our data suggests that if the N170 is sensitive to faces, this is limited to a small number of electrode sites, if any (Thierry et al., 2007a).

Importantly, this variability suggests differences in the topographies of the N170 to faces and cars. Although topographies of the N170 to different stimulus classes have been compared, using a variety of procedures like distribution comparisons (McCarthy & Wood, 1985), segmentation or microstate analyses, the pattern of results is inconsistent (Itier, Taylor, & Lobaugh, 2004; Thierry, et al., 2006). Regardless of this inconsistency, the validity of these findings is questionable. All topographic comparisons in general require the exclusion of amplitude differences; but because the topographies compared were confounded with the baseline topography, the necessary scaling will distort the distributions (Urbach & Kutas, 2002). In the present study, the face and car topographies in the upright full-front view conditions were of similar magnitude, enabling a valid distribution comparison without the need of scaling. We thus show for the first time that the topographies of the N170 for faces and cars are distinct. Distinct topographies are usually taken in favour of differences in the neural generation and thus in the underlying psychological processing (McCarthy & Wood, 1985; Picton et al., 2000). These functional differences were further supported by differences in the locations of the sources generating these potentials; the generators of the N170 for faces were located more inferior than the sources of the car N170, and the face and car dipoles in the left hemisphere had opposite orientations. Our finding thus suggests category sensitivity in the topographies of N170 to faces and cars.

Inversion reduced performance for faces, but increased performance for cars. Accuracy for detecting inverted compared to upright faces was reduced and accompanied by prolonged reaction times, whereas accuracy for detecting inverted compared to upright cars was increased, accompanied by a reduction in reaction times. This performance reduction for faces by inversion replicates the commonly known face inversion effect (Yin, 1969), affecting configural processing (Maurer et al., 2002). Interestingly, inverted cars were easier to classify as cars than upright cars. Future research will reveal whether this reflects easier perception of inverted cars as cars, or easier rejection of inverted cars as non-faces. The functional characteristics of inversion for cars are unknown and may be different from the mechanisms leading to the reduction in configural processing of faces.

Inverting faces delayed the N170 and increased its peak amplitude (Bentin et al., 1996). The N170 to inverted cars was similarly delayed, although to a smaller extent; this corroborates that the delay of the N170 is not unique to faces, but appears for many inverted stimuli (Itier et al., 2006; Rossion et al., 2003). In contrast to the increased amplitude to faces, inverting cars only negligibly modulated or reduced N170 amplitude. Taken by itself, this result supports the consensus that the increased amplitude by inversion is face sensitive (Bentin et al., 2007). On the other hand, inversion effects for faces have been lacking in instances, and other stimuli like words or houses have shown similar inversion effects (Eimer, 2000a; Itier & Taylor, 2004a; Rossion et al., 2000, 2003). These results challenge the consensus that inversion increases N170 amplitude reliably and solely to faces.

The direction of the inversion effect (i.e., the relative increase or decrease of amplitude between inverted and upright conditions) on N170 amplitude, where inversion effects were present, was opposite between faces and cars. The reduced amplitude for inverted cars may be a neural correlate of the increased performance for inverted cars. The difficulty in perceiving inverted faces has been associated with increased N170 amplitude (George, Evans, Fiori, Davidoff, & Renault, 1996; Rossion et al., 1999); our findings suggest that the effect of difficulty extends to non-face objects and can actually reduce N170 amplitude for stimuli that are easier to process when inverted.

The main aim of this study was to investigate whether presentation in full-front view abolishes the face sensitivity of the inversion effect in N170, that is, if full-front view cars and faces show a similar inversion effect. This was motivated by a similarly abolishing effect of full-front view presentation on the face sensitivity in N170 amplitude to upright objects (Thierry et al., 2007a). Here we showed that in fact stimulus variability also modulated the inversion effect, leading to larger amplitude changes by inversion for full-front view stimuli. In contrast, full-front view faces showed a large increase and full-front view cars a small reduction of N170 amplitude by inversion; thus, even after selecting comparable full-front view stimuli of similar size and symmetry, faces and cars show different inversion effects. Hence, full-front view presentation does not abolish the face sensitivity in the inversion effect on N170 amplitude. This finding suggests that inversion of faces indeed impairs face-sensitive processes like configural processing – first order relations, holistic processing and second order relations (Maurer et al., 2002), – and that this can be reliably measured in the N170.

Clear category differences in peak amplitude were present in P2. Faces elicited a larger P2 than cars. ERPs after the N170, for example the P2, have been linked to face-sensitive processing (Boehm, et al., 2006; Milivojevic, et al., 2003; Nasr & Esteky, 2009), and together these studies suggest later ERPs beyond N170 like the P2 are promising candidates for investigating perceptual processes sensitive to faces.

In sum, the N170 for upright stimuli was stronger for faces than cars only at the most inferior temporal sites, but of comparable amplitude at other sites, including some usually used for investigations of the N170 for faces. In addition, the N170 to faces and cars have distinct topographies. Inversion resulted in the commonly found increase in N170 amplitude for faces, but inversion of cars resulted only in negligibly modulated or reduced N170 amplitudes. The opposite direction of the inversion effect between categories echoed the increased performance for inverted cars and reduced performance for inverted faces. Stimulus variability modulated the effect of inversion on N170 amplitude for both categories, but using full-front view stimuli did not abolish the opposite and stronger modulation by inversion for faces than cars.

In conclusion, although the N170 was larger to faces than cars at the most inferior temporal sites, cars and faces can elicit similarly large amplitudes at other electrode sites commonly used for analyzing N170 to faces. This again calls into question the rationale that the N170 is face sensitive simply because of its larger amplitude to faces than any other category (Dering et al., 2009; Thierry et al., 2007a). Although the consensus that the increased N170 amplitude by inversion is face sensitive is not conclusively upheld by the overall pattern of results in the literature, it finds support in the present results. In contrast to amplitude differences between upright faces and other object categories, our results suggest that if the N170 is face sensitive, it is more likely reflected by its topography and the increased amplitude by inversion. This possible face sensitivity likely results from configural processing, which may affect the N170.

Chapter 4

Face-sensitive processes one hundred milliseconds after picture onset

This chapter is published as: Dering, B., Martin, C. D., Moro, S., Pegna, A. J., & Thierry, G. (2011). Face-sensitive processes one hundred milliseconds after picture onset *Frontiers in Human Neuroscience*, 5:93

Abstract

The human face is the most studied object category in visual neuroscience. In a quest for markers of face processing, event-related potential (ERP) studies have debated whether two peaks of activity –P1 and N170– are category-selective. Whilst most studies have used photographs of unaltered images of faces, others have used cropped faces in an attempt to reduce the influence of features surrounding the "face-object" sensu stricto. However, results from studies comparing cropped faces with unaltered objects from other categories are inconsistent with results from studies comparing whole faces and objects. Here, we recorded ERPs elicited by full-front views of faces and cars, either unaltered or cropped. We found that cropping artificially enhanced the N170 whereas it did not significantly modulate P1. In a second experiment, we compared faces and butterflies, either unaltered or cropped, matched for size and luminance across conditions, and within a narrow contrast bracket. Results of experiment 2 replicated the main findings of experiment 1. We then used face-car morphs in a third experiment to manipulate the perceived face-likeness of stimuli (100% face, 70% face and 30% car, 30% face and 70% car, or 100% car) and the N170 failed to differentiate between faces and cars. Critically, in all three experiments, P1 amplitude was modulated in a face-sensitive fashion independent of cropping or morphing. Therefore, P1 is a reliable event sensitive to face processing as early as 100 ms after picture onset.

Introduction

The human face is probably the most biologically significant stimulus encountered by humans in the environment because it provides critical information about other individuals (e.g., identity, age, sex, mood, direction of attention, intention, etc.). One fundamental question in visual neuroscience is whether or not the human ability to process face information relies on specific neural mechanisms qualitatively distinct from those involved in the perception of other classes of visual stimuli. A number of event-related potential (ERP) and magnetoencephalography (MEG) studies have been carried out to determine the time-course of category-selective effects during visual object perception and recognition. A particular peak of ERPs, the N170, which has a latency of ~170 ms after stimulus onset and is characterised by a vertex positive and bilateral temporal negative deflection (Bentin, Allison, Puce, Perez, & McCarthy, 1996; Grill-Spector & Malach, 2001; Linkenkaer-Hansen, et al., 1998), and its magnetic equivalent, the M170 (Liu, Harris, & Kanwisher, 2002; Xu, Liu, & Kanwisher, 2005), have been frequently reported as face-selective in the literature. In particular, it has been claimed that no stimulus category other than the human face elicits negativities as pronounced as faces in the 140-180 ms time-range after stimulus presentation (Itier & Taylor, 2004).

On the other hand, the P1, a peak with a latency of 100 ms, has also been suggested as a category-sensitive peak, albeit by a minority of authors (Herrmann, Ehlis, Ellgring, & Fallgatter, 2005; Thierry, Martin, Downing, & Pegna, 2007a). Despite the fact that P1 category-sensitivity has been repeatedly challenged (Bentin, et al., 2007; Kuefner, de Heering, Jacques, Palmero-Soler, & Rossion, 2010; Rossion & Jacques, 2008; but see also Dering, Martin, & Thierry, 2009; Thierry, Martin, Downing, & Pegna, 2007b), converging evidence from MEG, ERP and transcranial magnetic stimulation (TMS) have highlighted face sensitive processes occurring around 100 ms post-stimulus onset (Bentin & Golland, 2002; Herrmann, Ehlis, Ellgring, et al., 2005; Liu, Harris, & Kanwisher, 2002; Pitcher, Walsh, Yovel, & Duchaine, 2007). In particular, double TMS pulses have been shown to disrupt visual processing selectively for faces when stimulation is delivered over the Occipital Face Area (OFA) 60 and 100 ms after picture presentation but no measurable disruption is observed for double TMS pulses applied at later latencies (Pitcher, et al., 2007), nor when applied to nearby extrastriate areas.

Studies of intracranial recordings in patients with implanted electrodes have also yielded inconsistent results. Whist face-selective responses from the inferior temporal lobe have been recorded within 200 ms of stimulus onset (Allison, et al., 1994; Allison, Puce, Spencer, & McCarthy, 1999), other studies have suggested face-sensitive responses as early as 50 ms after stimulus onset (Seeck, et al., 2001; Seeck, et al., 1997), similar to some ERP studies (Braeutigam, Bailey, & Swithenby, 2001; Mouchetant-Rostaing & Giard, 2003). However, cortical activity in pharmacoresistant epileptic individuals can be affected by cognitive impairment after repeated seizures, anticonvulsant medication consumption, or functional reorganization subsequent to the presence of epileptic foci, making comparisons of intracranial recordings to ERPs only speculative (Allison, et al., 1999; Bennett, 1992; Krolak-Salmon, Henaff, Vighetto, Bertrand, & Mauguiere, 2004; Liu, et al., 2002).

A number of ERP studies have measured the sensitivity of the N170 peak to various stimulus manipulations in an attempt to determine which stage(s) of visual structural encoding are functionally reflected by the modulation of its amplitude. For instance, the N170 is sensitive to vertical orientation (Bentin, et al., 1996), isolation of internal features (Bentin, et al., 1996), scrambled facial features (George, Evans, Fiori, Davidoff, & Renault, 1996) as well as contrast (Itier & Taylor, 2002), spatial frequency (Goffaux, Gauthier, & Rossion, 2003) and gaussian noise (Jemel, et al., 2003). Since the N170 component is affected by the lack of internal (eyes, nose, mouth) and external (hair, ears, neck) features, it is likely to reflect –at least in part– configurational analysis of visual objects (Eimer, 2000b). Surprisingly, the sensitivity of the N170 to the external integrity of faces has rarely been investigated. Moreover, many studies of visual object categorisation have compared face and object perception using cropped faces (i.e., faces without hair, ears, or neck) and “intact” objects (Goffaux, et al., 2003; Jacques & Rossion, 2007; Kovacs, et al., 2006; Righart & de Gelder, 2007; Rossion, et al., 2003; Rousselet, Husk, Bennett, & Sekuler, 2007; Vuilleumier & Pourtois, 2007). Therefore, it is unclear whether differences between experimental conditions found earlier are indeed driven by categorical differences or artificially influenced by differences between experimental conditions in terms of stimulus integrity.

Why should stimulus integrity modulate N170 amplitude? When eyes presented in isolation are compared to pictures of complete faces, N170 amplitude is equally large, which has led to the hypothesis that the N170 may in fact index the activity of an eye-detection system (Eimer, 1998).

Furthermore, the N170 is highly sensitive to stimulus interpretability. That is, the same object can elicit larger N170 amplitudes when interpreted as part of a face (e.g., two dots interpreted as dots or as eyes; (Bentin & Golland, 2002).

Overall, because cropped faces and unaltered faces have often been used without distinction (Anaki & Bentin, 2009; Bentin, et al., 2007; Rossion, 2008; Rossion & Jacques, 2008; Zhao & Bentin, 2008), it is unknown whether face / non-face categorization takes place as early as 100 ms (P1 range) or beyond (N1 range). More specifically, a review paper by Rossion and Jacques (2008) has reported unpublished data as evidence against the findings of Thierry et al. (2007a). These results were based on faces and cars presented full front and repeated 6 times. Critically, the pictures of faces used were cropped but the pictures of cars were unaltered. Here, we investigated the effect of stimulus cropping and repetition to account for the discrepancies between the results obtained by Rossion and Jacques (2008) and those of Thierry et al. (2007a).

We presented participants with a stream of pictures featuring faces and cars (full front, symmetrical, centered, and of similar size within each condition). In the first block, all the stimuli were presented once complete and once cropped, i.e., faces without hair, ears, or neck and cars without rooftop, rear-view mirrors, or wheels (Figure 1A). In a second block, all stimuli used were repeated six times in order to test for potential repetition effects, since repetition is a factor inherent to previous studies on face categorization (e.g., Rossion and Jacques, 2008). This resulted in a within-participants 2 × 2 × 2 factorial design (face/car vs. cropped/unaltered vs. repetition/no repetition).

Participants performed a forced binary choice categorization task. We predicted (a) an effect of cropping onN170 amplitude for faces, which would account for the discrepancy between results obtained by Rossion and Jacques (2008) and Thierry et al. (2007a), without a category effect for the comparison of unaltered faces and cars; (b) a significant category effect on P1 amplitude replicating previous results (Thierry et al., 2007a; Dering et al., 2009; Boehm et al., 2011); and (c) an increase in P1 amplitude, and/or delayed P1 latency, by cropping, since categorization is arguably more difficult when peripheral information is missing, and given that we previously observed P1 amplitude increase with task difficulty (Dering et al.,2009).

In a second experiment, we sought to discard the hypothesis that effects of cropping on P1 or N170 amplitude could be due to residual differences in stimulus size, luminance, or contrast between experimental conditions by matching pictures with regard to all of these characteristics. We took this opportunity to compare the processing of faces to that of a third category, butterflies, which have been investigated previously (Schweinberger et al., 2004; Thierry et al., 2007a). The second experiment therefore had a 2 (faces/butterflies) ×2 (cropped/unaltered) design and featured no significant difference in size, luminance, between experimental conditions while keeping contrast variance within a narrow bracket[4] (Figure 1B). The predictions for Experiment 2 were exactly the same as that for Experiment 1.

---

[4] Due to cropping eliminating high contrast parts of the face–object (e.g.,hair), it was not possible to fully control for contrast across experimental conditions. However, significant effects of cropping – where they existed – always were in the opposite direction as that that would be expected from a contrast manipulation.

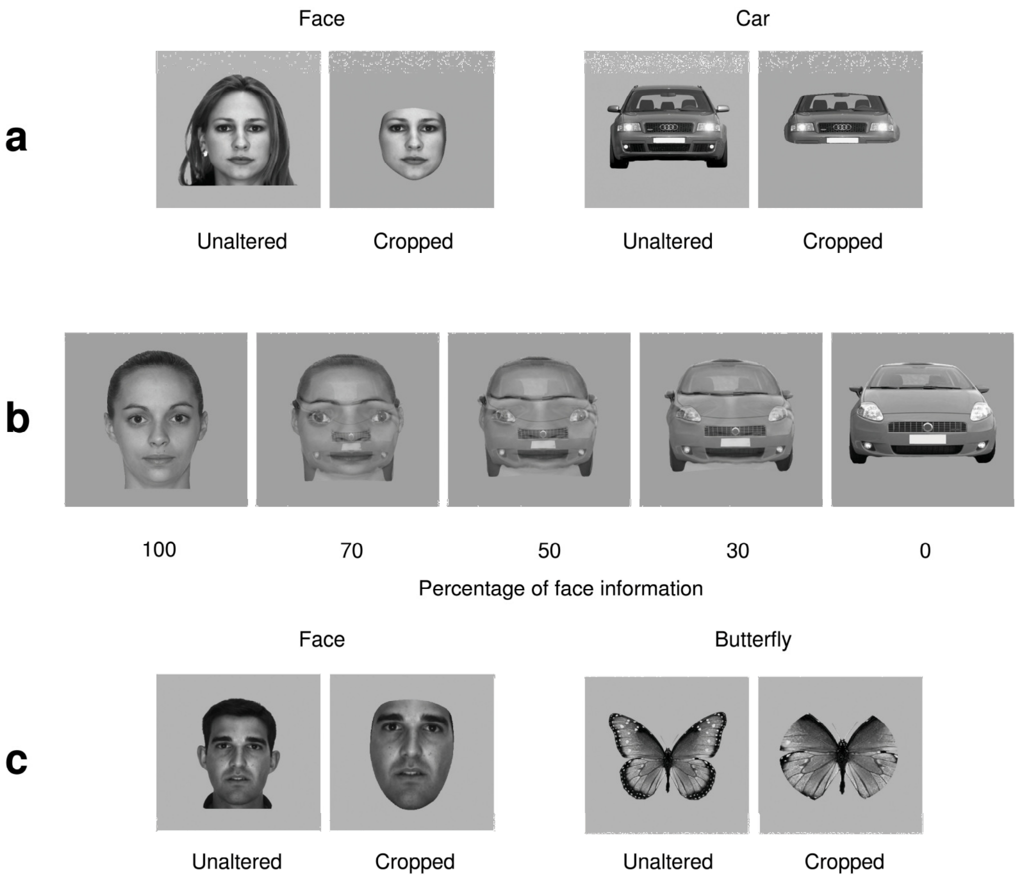

Figure 1 | **Examples of stimuli used in all experiments. (A) Example of a cropped face stimulus of the kind often used in experiments testing object categorization presented next to the unaltered source face and a similar example comparison for a car stimulus. (B) Examples of cropped and unaltered faces and butterflies after matching for luminance, contrast, and size. (C) Examples of progressive morphing between a face and a car stimulus. Note that 50% face–50% car morphs were used as target stimuli and not analyzed.**

In a third experiment, we manipulated stimulus interpretability. Full front views of faces and cars were morphed to produce images that contained face and car information in various proportions: 100% face, 70% face–30% car, 50% face–50% car, 30% face–70% car, and 100% car (Figure 1C). The ambiguous 50% face–50% car condition was highlighted by a frame, required participant's responses, and was discarded at the analysis stage. This resulted in a 2 (face vs. car) ×2 (morphed vs. unaltered) design. Any component that is presumed to be face-sensitive was predicted to be significantly larger for face-like stimuli as compared to car-like stimuli.

Materials and Methods

Experiment 1

Twenty-two participants (mean age = 24.5, SD = 5.5, 15 females, 1 left-handed) with normal or corrected-to-normal vision gave written informed consent to participate in the experiment that was approved by the ethics committee of Bangor University. Ninety-six images of full front faces were modified digitally so as to remove features considered peripheral to the face-object *sensu stricto*, i.e., hair, ears, and neck. Ninety-six images of full front cars were modified in a similar way by removing roof top, wing mirrors and wheels. After cropping, all images were transposed onto a grey background (Fig 1). All images in each of the 4 groups generated (cars and faces, unaltered and cropped) were centred on the screen, scaled to fit a standard size template, and had the same orientation (thus reducing stimulus variability as much as possible (Thierry, et al., 2007a). Cropping images resulted in slight variations in luminance (cropped faces 42.1 cd/m^2; cropped cars 40.1 cd/m^2; unaltered faces 34.9 cd/m^2; unaltered cars 39.8 cd/m^2) and contrast (cropped faces 0.7 cd/m^2; cropped cars 1.5 cd/m^2; unaltered faces 3.6 cd/m^2; unaltered cars 3.5 cd/m^2) between conditions. In a first variant of the experiment, stimuli were presented in a randomized order in 4 blocks of 96 trials such that each block featured 24 pictures from each of the four experimental conditions. In a second variant of experiment 1, a selection of 16 images repeated 6 times each were presented in the same randomized fashion to test for potential effect of repetition. Stimuli were presented for 200 ms, with an inter-stimulus interval of 1300 ms, and participants categorised each of the stimuli as face or car by pressing keys on a keyboard, a task shown to elicit similar ERP patterns as a one-back task (Dering, et al., 2009). Response sides were counterbalanced between participants.

Experiment 2

Twenty participants (mean age = 22.1, SD = 3.7, 11 females, 1 left-handed) with normal or corrected-to-normal vision gave written informed consent to participate in the experiment that was approved by the ethics committee of Bangor University. Eighty images of full front faces were modified as in experiment 1, by removing features considered peripheral to the face-object. Eighty images of full front butterflies were modified in a similar way by cropping the wings. After cropping, the images were enlarged without distortion to fit the maximum x and y dimensions of their counterpart unaltered image, and then transposed onto a grey background (Fig 1). All images in each of the 4 groups generated (butterflies and faces, unaltered or cropped) were centred on the screen, matched for size, and had the same orientation (thus reducing stimulus variability as much as possible (Thierry, et al., 2007a). Relative luminance was set at 42 cd/m^2, and contrast was 17.8 cd/m^2 ±2.2 on average including the grey background in the calculation. Participants sat 100 cm from a calibrated CRT monitor. Stimuli were presented for 200 ms, with an inter-stimulus interval of 1300 ms, and participants categorised each of the stimuli as a face or butterfly by pressing keys on a stimulus response box. Response sides were counterbalanced between participants. Stimuli were presented in a randomized order in 4 blocks of 160 trials such that all images in the experiment were presented twice.

Experiment 3

Eighteen participants (mean age = 19.8, SD = 1.99, 13 females, 0 left-handed) with normal or corrected-to-normal vision gave written consent to participate in the experiment that was approved by the ethics committee of Bangor University. Forty images of full front neutral faces aged between 18 – 30 years old were obtained from the Productive Aging lab's face database (Minear & Park, 2004). These images, centred on the screen, scaled to fit a standard size template, and with the same orientation were transposed onto a uniform grey background. Forty pictures of full front faces were paired with 40 pictures of full front cars and transformed using a morphing algorithm (Sqirlz Morph 2.0) to produce a series of face-car morphs varying in the percentage of face information embedded in each image (Fig. 1): 100% face, 70% face–30% car, 50% face–50 % car, 30% face–70% car, and 100% car. Stimuli were presented for 500 ms in a randomised order followed by an interval of 1500 ms allowing for participant response. The morphing procedure produced slight variations in luminance such that across all conditions average luminance was 37.2 cd/m^2, ± 1 cd/m^2 (contrast 1.5 cd/m^2). Each picture was presented 6 times throughout the experiment. Participants only responded to ambiguous target stimuli (50% face–50 % car), which were presented in a distinctive black frame, by indicating whether the picture was perceived rather as a face or as a car. The task was a forced-choice binary task and response sides were counterbalanced between participants. All stimuli across the three experiments subtended no more than 8.53^0 of horizontal and vertical visual angle.

Event-related potentials

Using Cz as a reference, scalp activity was recorded using SynAmps2™ (Neuroscan, Inc., El Paso, TX, USA) amplifiers with a sampling rate of 1 kHz from 64 Ag/AgCl electrodes (Easycap™, Brain Products, Germany) distributed across the scalp according to the extended 10–20 system. Impedances were kept below 5 kΩ. The electroencephalogram was filtered on-line between 0.01 and 200Hz and off-line with a low-pass zero phase shift digital filter set to 30Hz (48 db/octave slope). Eye-blink artifacts were mathematically corrected[5] using a model blink artifact computed for each individual following the procedure recommended by Gratton et al. (1983). Signals exceeding ±75µV in any given epoch were automatically discarded. EEG recordings were cut into epochs ranging from −100 to 500 ms after stimulus onset and averaged for each individual in all experiments according to the experimental conditions. Grand-averages were calculated after re-referencing individual ERPs to the common average reference. Mean amplitudes for each condition were analyzed at eight posterior occipital

---

[5] To check the efficacy of our eye-blink correction procedure and to establish that residual noise did not affect our results, we conducted a new analysis excluding all trials containing eye movement artifacts. The net loss of trials was <1.4% and statistical results were unchanged in all three experiments.

electrodes for Experiment 1. Global field power was calculated to guide classification of ERP components (Koenig & Melie-Garcia, 2010). Peak latencies were measured at the electrode of maximal amplitude in each condition and each participant. The P1 was identified as a positive peak occurring between 80 and120ms and analyzed at sites O1, O2, PO7, PO8, PO9, and PO10. Due to significant differences between latencies for conditions at the P1, mean amplitude analyses were conducted 20ms around the peak of maximal activity for each condition of the experiment (Table 1). The N170 peaked between 120 and 200ms at electrode sites P7, P8, PO7, PO8, PO9, and PO10. Mean amplitude analyses for the N170 were conducted 40ms around the peak for each condition of the experiment (Table 1).

Table 1 **Average P1 and N170 peak latencies (in milliseconds) in experiments 1, 2 and 3.**

Experiment 1					
		UF	**CF**	**UC**	**CC**
P1	*no stimulus repetition*	**93**	**101**	**96**	**105**
N170		**136**	**145**	**147**	**155**
P1	*stimulus repetition*	**93**	**100**	**93**	**104**
N170		**136**	**147**	**149**	**157**

Experiment 2				
	UF	**CF**	**UB**	**CB**
P1	**92**	**98**	**86**	**87**
N170	**143**	**146**	**143**	**147**

Experiment 3				
	100% Car	**70% Car**	**70% Face**	**100% Face**
P1	**109**	**104**	**105**	**102**
N170	**157**	**157**	**154**	**148**

The data was subjected to repeated measures analysis of variance (ANOVAs) with three factors – category (face/car), alteration (unaltered/cropped), and electrode (6 levels). A Greenhouse-Geisser correction was used where applicable. To demonstrate the magnitude of effects, partial Eta squared (ήp2) is reported. In the analyses reported here, the electrode factor was systematically significant but such effects are not discussed since the focus of this paper was on mean peak amplitude differences at electrodes of predicted (and observed) maximal sensitivity. For a contribution addressing the issue of topographical comparisons, see Boehm et al. (2011).

For experiments 2 & 3, P1 and N170 components peaked within the same time windows used for analysis in experiment 1, as indicated by calculation of the global field power. P1 and N170 were examined at the same electrode sites as experiment 1 respectively, with mean amplitude analyses for P1 run 20 ms around each peak for each condition and 40 ms around the peaks for N170 (Table 1). Experiment 2 was analysed by repeated measures ANOVAs with 3 factors of category (face/butterfly), cropping (unaltered/cropped) and electrode (6 levels). Experiment 3 had 3 factors of category (face/car), morphing (morphed images/normal images) and electrode (6 levels). Effect sizes (ήp2) are also reported where relevant.

Temporal segmentation

This analysis tracked scalp topographies that remain stable for periods of time in the order of tens to hundreds of milliseconds (Michel, et al., 2001). These so-called microstates are thought to represent specific phases of neural processing (Brandeis & Lehmann, 1986; Lehmann & Skrandies, 1984; Michel, Seeck, & Landis, 1999; Michel, et al., 2001). We identified the microstates using a hierarchical cluster analysis technique (Michel, et al., 2001; Pascual-Marqui, Michel, & Lehmann, 1995) to determine the segmented maps accounting for the greatest amount of variance in the ERP map series. The optimal number of segment maps explaining the greatest amount of variance was obtained using a cross-validation criterion (Michel, et al., 2001; Murray, Brunet, & Michel, 2008; Pegna, Khateb, Michel, & Landis, 2004; Pegna, et al., 1997; Thierry, et al., 2007a; Thierry, et al., 2006; Vuilleumier & Pourtois, 2007). Then, we calculated the statistical validity of maps extracted from grand-averages by determining the amount of variance explained by each map in the ERPs of each individual in each condition. Repeated measures ANOVAs were then performed on these values to determine the quality of fit between microstates and individual data (Murray, et al., 2008; Pegna, et al., 2004; Pegna, et al., 1997; Thierry, et al., 2007a; Thierry, et al., 2006).

Results

Cropping faces artificially increases N170 amplitude but does not affect P1 category-sensitivity

In experiment 1, the mean reaction time was 381 ms ±77 across all conditions and mean accuracy was 92 ±6.6%. Neither reaction times nor accuracy was affected by stimulus category or cropping (all Ps > 0.05).

ERPs for all 22 participants displayed a typical P1-N1-P2 complex in all experimental conditions (Fig. 2). Analysis of P1 amplitudes revealed a pattern of response sensitive to face information present in the stimulus. Repeated measures analysis of variance (ANOVA) over 6 posterior occipital electrodes revealed no main effect of repetition on P1 mean amplitude [F(1,21)=2.482, p>.05] allowing data for the unrepeated and repeated blocks to be combined for further analysis. There was a main effect of object category on P1 mean amplitudes [F(1,21)=15.87, p<.05, ήp2 = 0.43] showing that the P1 elicited by faces was significantly larger than the P1 elicited by cars but, there was no effect of cropping on P1 mean amplitude [F(1,21)=2.507, p>.05] and critically, no interaction between the two factors [F(1,21)=0.621, p>.05]. Conversely, P1 peak latency was significantly delayed by cropping [F(1,21)=15.4, p<.05, ήp2 = 0.423] but no other experimental factors (all ps> .1).

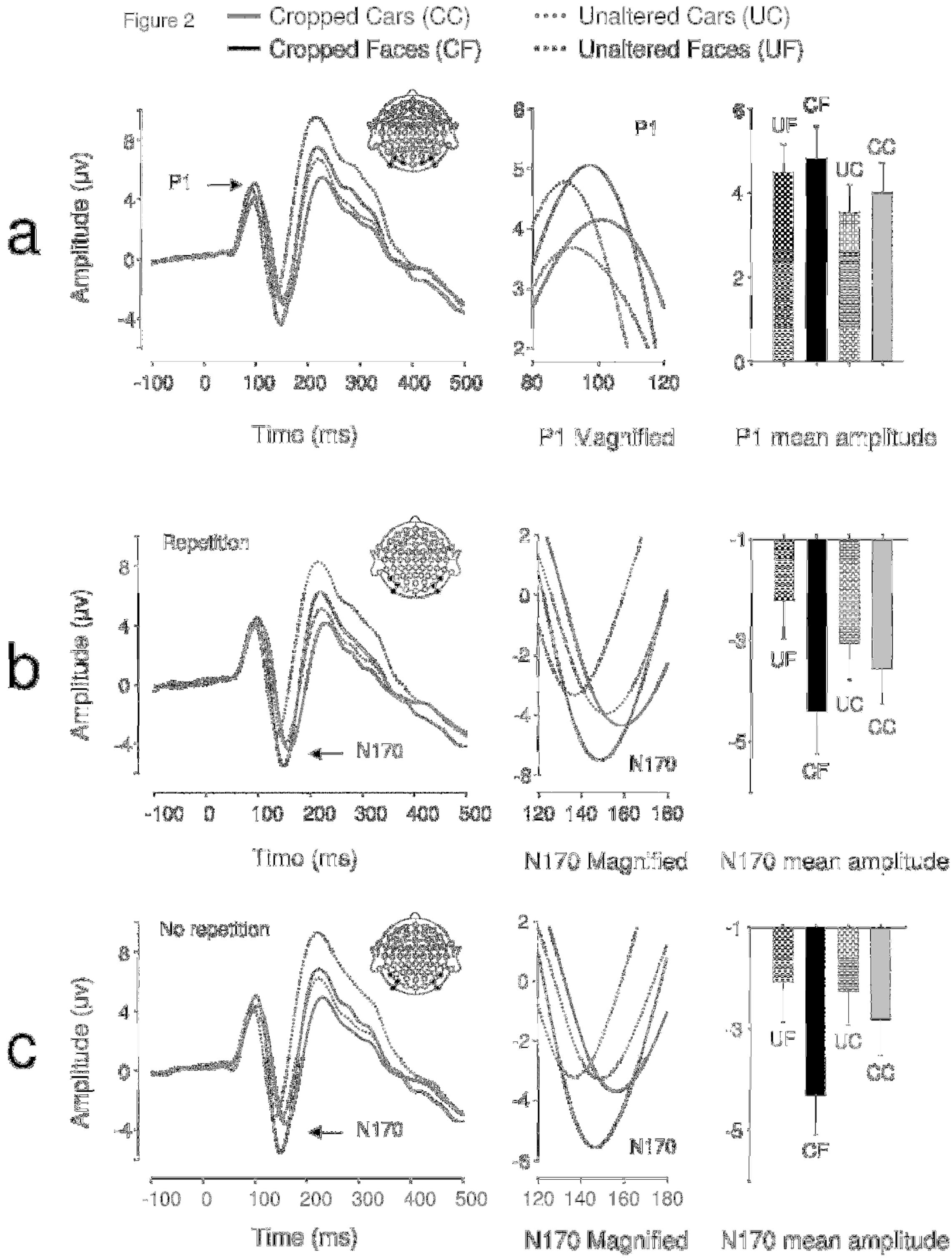

Figure 2 – **Grand averaged event-related brain potentials recorded in the four conditions of Experiment 1. Waveforms depict a linear derivation of the electrodes used in the statistical analysis for the P1 and N170, respectively. (a) From left to right:**

Linear derivation of electrodes O1, O2, PO7, PO8, PO9 and PO10 regardless of stimulus repetition (factor non significant), magnification of the P1, and bar plot of P1 mean amplitudes. (b) From left to right: Linear derivation of electrodes P7, P8, PO7, PO8, PO9 and PO10 in the experimental block featuring stimulus repetitions, magnification of the N170, and N170 mean amplitudes. (c) Linear derivation of electrodes P7, P8, PO7, PO8, PO9 and PO10 in the experimental block without stimulus repetition, magnification of the N170, and N170 mean amplitudes. Error bars depict s.e.m.

We found a main effect of repetition on N170 mean amplitude [F(1,21)=7.13, p<.05, ήp2 = 0.253]. Stimulus repetition increased N170 amplitude and this effect was greater for cars than faces as indicated by a significant repetition by category interaction [F(1,21)=14.04, p<.05, ήp2 = 0.401]. The repetition factor did not interact with any other factors. As predicted, object category failed to modulate N170 mean amplitude [F(1,21)=0.799, p>.1] (Fig. 2b). However, there was a main effect of cropping [F(1,21)=43.001, p<.0001, ήp2 = 0.672], such that cropped stimuli elicited greater N170 mean amplitudes than unaltered stimuli. Also, cropping and object category interacted [F(1,21)=43.37, p<.0001, ήp2 = 0.675], showing that the difference in N170 mean amplitude between the cropped and unaltered conditions was greater for faces than cars (Fig. 2b). Bonferroni-corrected pair-wise comparisons between cropped and unaltered objects were significant both in the case of faces (p<.0001) and in that of cars (p<.001). It is noteworthy that the contrast often reported in the literature, i.e., cropped faces vs. unaltered cars was highly significant (p<.0001).

There was no main effect of repetition on N170 latencies [F(1,21)=0.54, p>.01]. Repetition interacted with category [F(1,21)=8.401, p<.05, ήp2 =

0.286], reducing latencies for repeated in comparison to unrepeated cars [F(1,21)=6.53, p<.05, ήp2 = 0.237], however this latency difference was only 2 ms. Repetition did not interact with any other factor [Fs(1,21)<=1.85, p>=.188]. N170 latency was significantly modulated by both cropping [F(1,21)=255.72, p<.05, ήp2 = 0.924, delayed for cropped as compared to unaltered stimuli] and category [F(1,21)=123.82, p<.05, ήp2 = 0.855, delayed for cars as compared to faces], and these factors interacted significantly [F(1,21)=7.5, p<.05, ήp2 = 0.263], showing that cropping had a greater influence on N170 latencies for faces than cars.

In order to determine the sensitivity of the P1 to object category and that of the N170 to cropping in experiment 1, a segmentation analysis was performed on the map series elicited in each condition between 0 and 250 ms. This procedure identified two distinct maps for P1 (P1a and P1b) and two distinct maps for N170 (N1a and N1b). The statistical validity of the microstates was tested by evaluating the amount of variance explained by the maps issued from segmentation in each individual participant maps in each condition (see methods) using a 2 (cropping) x 2 (category) x 2 (maps) repeated measures ANOVA (Fig.3a).

For unrepeated and repeated blocks separately, category by map interactions [Fs(1,21)>15.5, ps<.001] and univariate tests for planned comparisons confirmed that P1a explained a significantly greater proportion of variance for faces than cars [Fs(1,21)>11.4, ps<.01]. Conversely, map P1b better explained individual maps for cars than faces in the unrepeated block only [F(1,21)=12.0, p<.001]. We found no effect of cropping on P1 microstates (all Ps>.05). In sum, as predicted, microstates in the P1 range patterned with category differences rather than repetition or cropping.

Figure 3

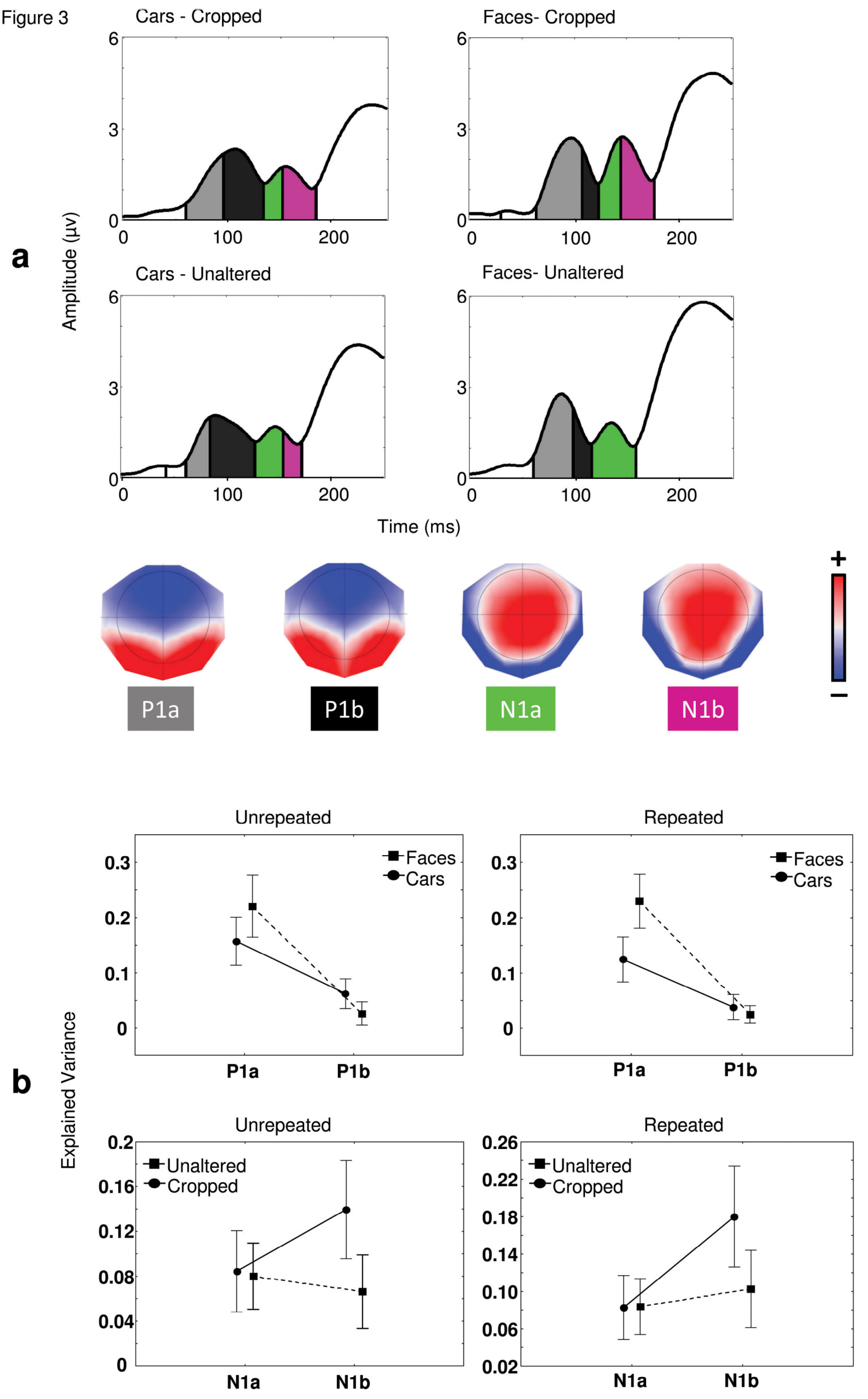
a
Cars - Cropped
Faces- Cropped
Cars - Unaltered
Faces- Unaltered
Amplitude (µv)
6
3
0
0
100
200
Time (ms)
+
−
P1a
P1b
N1a
N1b
b
Explained Variance
Unrepeated
Repeated
Faces
Cars
0.3
0.2
0.1
0
P1a
P1b
Unaltered
Cropped
0.2
0.16
0.12
0.08
0.04
0
0.26
0.22
0.18
0.14
0.10
0.06
0.02
N1a
N1b

Figure 3 - **(a) Global field power waveforms in the 4 conditions of experiment 1 segmented by microstate and associated topographies identified by the segmentation procedure (see methods). (b) proportion of explained variance by maps P1a and P1b regarding object categories and maps N1a and N1b in relation to cropped and unaltered images.**

In the N1 range, in both the case of repeated and unrepeated blocks, we found significantly different microstates between cropping conditions [Fs(1,21)>6.1, ps<.05], such that map N1b better explained variance for cropped than unaltered conditions, whereas map N1a failed to distinguish between any of the experimental conditions [Ps>.1]. Univariate test for planned comparisons confirmed that cropping x map interactions were due to map N1B explaining individual maps for cropped stimuli significantly better than individual maps for unaltered stimuli [F(1, 21)=20.3, p<.001], while N1A produced no difference [F(1, 21)=.01, p>.05].

N170 cropping effects are not driven by low-level differences between experimental conditions

In experiment 2, the mean reaction time was 399 ms ±54 across all conditions and mean accuracy was 95 ±3.2 %. Accuracy was not affected by either stimulus category or cropping (all Ps > 0.05). Reaction times differed significantly for category [F(1,19)=5.3, p<.05, ήp2 = 0.22] and cropping [F(1,19)=11.57, p<.05, ήp2 = 0.38], and these factors interacted [F(1,19)=13.6, p<.05, ήp2 = 0.42]. Overall, reaction times were slower to faces than butterflies, with unaltered faces producing the largest delay.

ERPs for all 20 participants displayed a typical P1-N1-P2 complex in all experimental conditions (Fig. 4). We found a main effect of object category on P1 mean amplitudes [F(1,19)=6.29, p<.05, ήp2 = 0.25] such that faces elicited greater P1s than butterflies. Critically, with stimuli matched for size and luminance across conditions, there was no effect of cropping on P1 mean amplitude, as in experiment 1[F(1,19)=0, p>.05]. Furthermore, object category and cropping did not interact [F(1,19)=0.887, p>.05]. P1 peak latency was unaffected by object category but significantly delayed by cropping [F(1,19)=8.44, p<.05, ήp2 = 0.31] and the two factors interacted [F(1,19)=4.17, p<.05, ήp2 = 0.2] such that cropped faces delayed P1 latencies more than cropped butterflies (although this was a 2 ms difference).

In the N170 range, as expected from experiment 1, cropped stimuli elicited greater N170 mean amplitudes than unaltered stimuli [F(1,19)=21.61, p<.0001, ήp2 = 0.53] (Fig. 4b). But unexpectedly, faces elicited significantly greater N170 mean amplitudes overall [F(1,19)=17.12, p<.05, ήp2 = 0.47]. However, this effect was driven by cropping, as indicated by a significant category by cropping interaction [F(1,19)=40.82, p<.0001, ήp2 = 0.68], such that the difference in N170 mean amplitude between cropped and unaltered conditions was greater for faces than butterflies (Fig. 4b). Critically, unaltered faces and butterflies did not significantly differ in mean amplitude [F(1,19)=2.13, p>.1]. Finally, cropping increased N170 latencies [F(1,19)=23.8, p<.0001, ήp2 = 0.556] by 3 ms on average, but no other factor affected N170 latencies [Fs(1,19)<1.04, ps>.1].

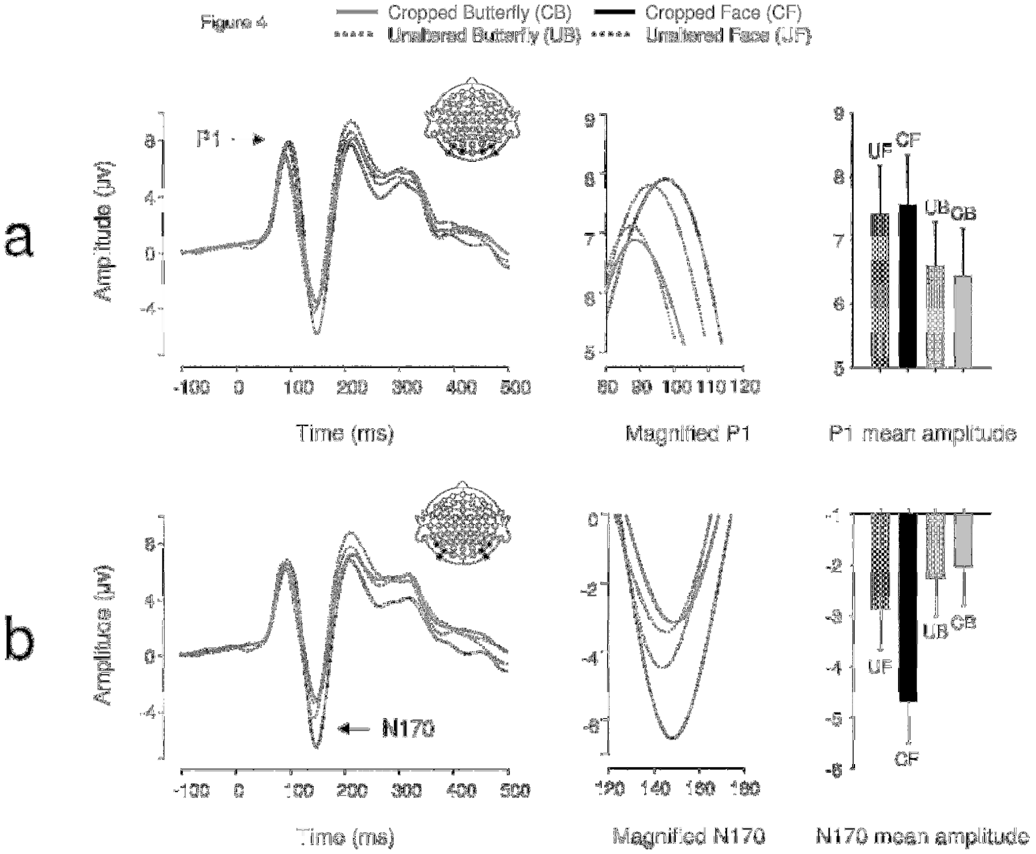

Figure 4 – **Event related brain potential results in the four conditions of Experiment 2. Waveforms depict linear derivations of the electrodes used for analysis of the P1 and N170, respectively. (a) From left to right: Linear derivation of electrodes O1, O2, PO7, PO8, PO9 and PO10, magnification of the P1, and bar plot of P1 mean amplitudes. (b) From left to right: Linear derivation of electrodes P7, P8, PO7, PO8, PO9 and PO10, magnification of the N170, and bar plot of N170 mean amplitudes. Error bars depict s.e.m.**

Morphing pictures across categories affects P1 but not N170 amplitude

In experiment 3, a repeated measures ANOVA performed over 6 posterior occipital electrodes revealed a main effect of category [F(1,17)=18.09, p<.005, ήp2 = 0.516] and morphing [F(1,17)=6.44, p<.05, ήp2 = 0.275] on P1 mean amplitude. There was no interaction between these factors [F(1,17)=1.9, p>.1], suggesting that category and morphing independently increased P1 amplitude (Fig 5). Previous findings of a categorical difference within the P1 range were confirmed [F(1,17)=11.92, p<.01, ήp2 = 0.412] with faces eliciting larger P1 amplitudes than cars. Furthermore, P1 mean amplitude was significantly greater for 100% face and 70% face (30% car) stimuli than 100% car and 70% car (30% face) stimuli, respectively (all ps<.05). In other words, P1 mean amplitude was systematically greater for face-like than car-like stimuli. Finally, no differences were found between the 70% face (30% car) and the 100% face conditions [F(1,17)=1.27, p>.1], but there was a difference between 100% car and 70% car (30% face) [F(1,17) = 7.98, p=.012, ήp2 = 0.32]. Neither morphing stimuli or categorical differences affected P1 latency [Fs(1,17)<1.244, ps>.1].

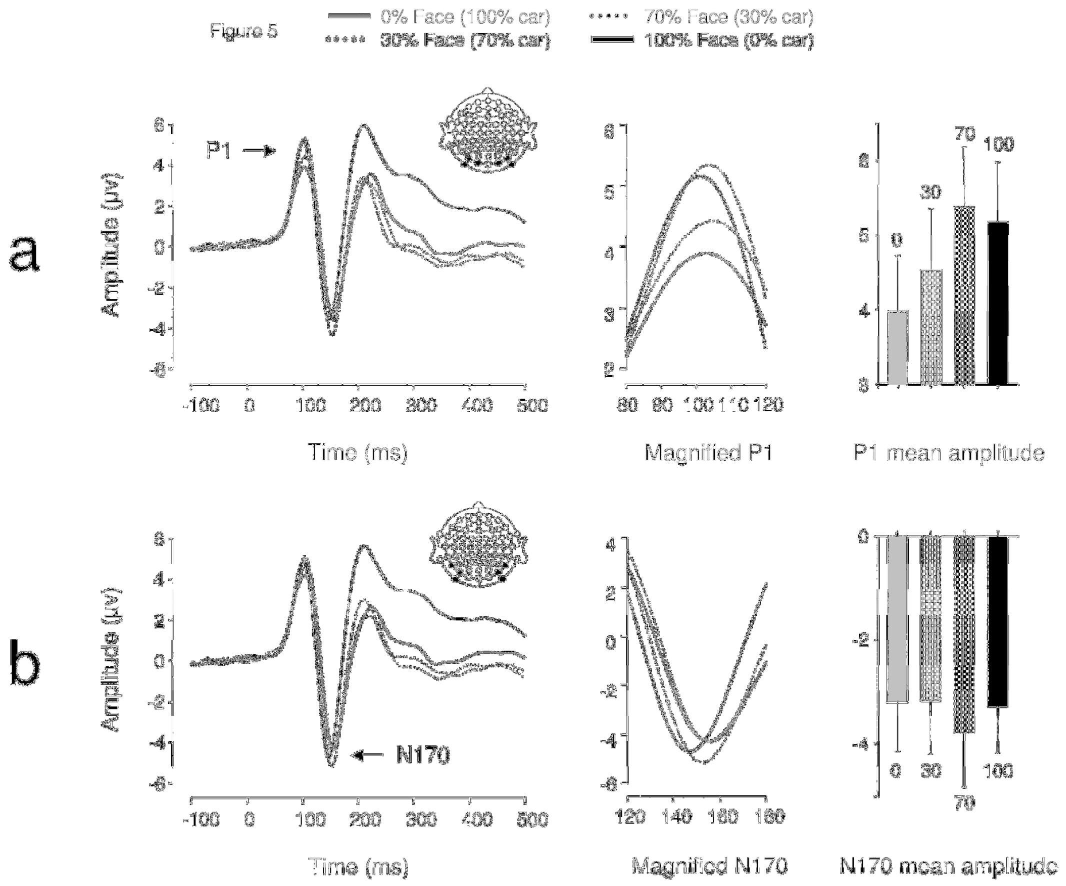

Figure 5 – **Event related brain potential results in the four conditions of Experiment 3. Waveforms depict linear derivations of the electrodes used for analysis of the P1 and N170, respectively. (a) From left to right: Linear derivation of electrodes O1, O2, PO7, PO8, PO9 and PO10, magnification of the P1, and bar plot of P1 mean amplitudes. (b) From left to right: Linear derivation of electrodes P7, P8, PO7, PO8, PO9 and PO10, magnification of the N170, and bar plot of N170 mean amplitudes. Error bars depict s.e.m.**

We found no significant modulations of amplitude or latency by either category or morphing in the N170 range [Fs(1,17)<2.881, ps>.1].

Discussion

N170

Experiment 1 aimed at testing the effect of discarding peripheral visual information when studying the neurophysiological indices of face processing. While various lines of evidence have challenged the face-selectivity of the N170, this component remains widely regarded as face-selective (Bentin, et al., 1996; Blau, et al., 2007; Carmel & Bentin, 2002; Eimer, et al., 2010; Itier & Taylor, 2002; Mohamed, et al., 2009; Rossion & Jacques, 2008; Sadeh, et al., 2008). In all three experiments reported here however, the N170 failed to behave in a face-selective manner (Thierry et al., 2007a, but see Bentin et al., 2007b; Thierry et al., 2007b; Rossion and Jacques, 2008).

Full front views of cars were compared to faces in Experiments 1 and 3 because the two categories have properties in common: they are highly frequent and familiar objects, easy to categorize, susceptible of being subcategorized (make/ethnic origin), they have generic internal features, the arrangement of which is critical for identification. We considered that cars are the ideal contrast category for faces precisely because of these shared properties since a brain response selective to faces should indeed distinguish between the two categories regardless of their similarities. In previous studies involving full front views of faces and cars, N170 selectivity was not measurable (Rossion et al., 2000; Schweinberger et al., 2004; Thierry et al., 2007a). A common account for the finding of similar amplitudes to faces and cars is that they are perceptually highly similar, which could evoke comparable N170 responses (Hadjikhani, Kveraga, Naik, & Ahlfors, 2009) but, ultimately, unless neuropsychologically impaired, no one could ever claim that a picture of a car can be confused with that of a face.

A number of studies have resorted to cropping faces from full head pictures, particularly for behavioral testing of face recognition in patients with prosopagnosia (Behrmann, Avidan, Marotta, & Kimchi, 2005; I. Gauthier, Behrmann, & Tarr, 1999; Herzmann, Kunina, Sommer, & Wilhelm, 2009; Saumier, Arguin, & Lassonde, 2001; Stollhoff, Jost, Elze, & Kennerknecht, 2010). The rationale behind the use of such modified stimuli in neuropsychological testing is to prevent patients relying on the analysis of peripheral cues such as hair color and shape, neck width, or ear size and shape to recognize faces (Duchaine & Nakayama, 2004). The fact that faces are more difficult to recognize when peripheral cues are removed implies that such cues are important in the natural process of face recognition. Surprisingly, however, experimental psychologists and neuroscientists have used such cropped faces as stimuli in experiments testing visual object categorization without preliminarily testing whether this alteration would affect the processing of different object categories in different ways (Horovitz, et al., 2004; Kuefner, de Heering, Jacques, Palmero-Soler, & Rossion, 2010; Schiltz, et al., 2006). Our findings show a dramatic effect of such stimulus alteration in the case of faces as compared to the case of non-face stimuli, the N170 being increased in amplitude by 1.37µV on average and 1.39µV at the peak when peripheral features are deleted. This effect is consistent with modulation of hemodynamic responses from the fusiform face area (FFA) and OFA found for external features of faces presented in isolation (Andrews, Davies-Thompson, Kingstone, & Young, 2010).

Interestingly, a significant change in N170 mean amplitude was even found in the case of cars, albeit of smaller amplitude. It is worth noting here that cropping cars is less straightforward than cropping faces, since the internal features of a face are easily identifiable whereas those of a car are uncertain. This is mainly due to the existence of a neutral, featureless area (e.g., forehead, cheeks, etc.,) between peripheral and inner parts in the case of faces that has no equivalent in the case of cars. In addition, spatial relations of features tend to differ between manufacturer's models of cars to a greater extent than between individual faces. Therefore, our cropped car stimuli were probably less representative of inner part extraction than cropped face stimuli, which may have accounted for the relatively smaller amplitude modulation by cropping for cars. It is noteworthy that the cropped versions of faces and cars were smaller in size as compared to unaltered images. It has been shown that ERP amplitudes increase with stimulus size (Busch, Debener, Kranczioch, Engel, & Herrmann, 2004; De Cesarei & Codispoti, 2006), therefore differences in physical size between cropped and unaltered stimuli are unlikely to account for the increase in amplitude by cropping observed in the N170 range. In Experiment 2, we addressed this issue directly by matching the cropped stimuli with the unaltered stimuli in terms of size (and luminance while keeping to a narrow contrast range). The effect of cropping on N170 amplitude for faces was fully replicated. However, we did not find such an effect for butterflies. This arguably is not surprising since cropped butterflies were perceptually similar to unaltered ones (see Figure 1 for an example). Since no significant information was lost by cropping (i.e., only a portion of the wings), there is no reason why visual processing should have been more difficult in this condition even though cropping obviously engendered ineluctable differences in spatial frequency between conditions.

In addition to a significant modulation by cropping in Experiment 1, we found a main effect of stimulus repetition on N170 mean amplitude. Unexpectedly, the repetition effect was an increase in amplitude with repetition, which was greater for faces than cars. This is inconsistent with the previously reported habituation effect (Campanella, et al., 2000; Heisz, Watter, & Shedden, 2006), which would be expected to result in an N170 amplitude reduction triggered by immediate repetition. However, repetitions in our study were always separated by several intervening trials, possibly supporting a familiarity account of N170 modulation (e.g., (Jemel, Schuller, & Goffaux, 2010; Leleu, et al., 2010; Tacikowski, Jednorog, Marchewka, & Nowicka, 2011). Indeed, in studies of repetition priming where long lags between repeated stimuli have been used, no reduction of N170 amplitude has been reported (Schweinberger et al., 2002; Boehm et al., 2006). Furthermore, larger amplitudes to repeated stimuli were previously reported in the N1 range when degraded images were used as primes (Doniger, et al., 2001).

We can draw two conclusions from the N170 findings: (a) Although the N170 was not category-selective when we compared complete faces, cars, and butterflies, it is strongly amplified when features important for face recognition are deleted, and we contend that N170 is likely to index mechanisms beyond object categorization such as the processing of familiarity, identity, ethnic origin, emotional expression, etc.; (b) the comparison of N170 amplitude elicited by cropped faces and other object categories presented without alteration should not be used to make claims regarding category-selectivity in visual cognition.

P1

We found events compatible with object categorization in the P1 range when comparing faces and cars (Experiment 1) or faces and butterflies (Experiment 2), and sensitivity to stimulus category was independent of other manipulations, e.g., stimulus variability (Thierry et al., 2007a), cropping (Experiments 1 and 2), and morphing (Experiment 3).

The P1, a peak generally regarded as an index of low-level perceptual processing (Picton et al., 2000; Tarkiainen et al., 2002; Cornelissen et al., 2003; Rossion et al., 2003) and repeatedly suggested as being sensitive to differences in contrast, color, luminance, etc. (Nakashima, et al., 2008; Thierry, Athanasopoulos, Wiggett, Dering, & Kuipers, 2009), was only sensitive to object category (i.e., to global image features) regardless of stimulus integrity. In Experiment 1, there were residual differences between cropped and unaltered stimuli in terms of luminance and contrast because peripheral features of faces, in particular, tend to have high contrast and low luminance and thus cannot be dismissed without affecting low-level properties of stimuli. Nevertheless, the small difference in luminance between cropped and unaltered stimuli in Experiment 1 should have produced a P1 modulation in the opposite direction to the trend observed (e.g., Thierry et al., 2009; and in any case the effect of cropping on P1 amplitude was not significant). Furthermore, in Experiment 2 in which stimuli were matched for luminance and narrowly controlled in terms of contrast, the effect of category on P1 amplitude was fully replicated.

Overall, category-sensitivity in the P1 range held across all three experiments and was consistent with findings of a critical phase of visual object categorization at around 100ms post-stimulus presentation in MEG, ERP, and TMS studies (Liu et al., 2002; Herrmann et al., 2005; Pitcher et al., 2007). In Experiment 3, the P1 was not only increased for faces relative to cars but also for stimuli affected by morphing as compared to unaltered stimuli. This result is consistent with the view that P1 amplitude is increased by visual ambiguity (Schupp, et al., 2008) and, more generally, task difficulty (Dering et al., 2009) because categorization of morphed stimuli – which contain information from the other category– is more challenging than that of unaltered images.

Conclusion

To our knowledge, the effect of cropping inner parts of faces and objects on visual categorization has never been studied directly using ERPs and the potential effects of this manipulation have not been discussed (Eimer, 2000b; Duchaine and Nakayama, 2004). This leads to the possibility that category-effects previously reported in the N170 range may have been due not only to uncontrolled perceptual variance between conditions (Thierry et al., 2007a) but also reduction in the amount of information afforded by artificially impoverished stimuli. Furthermore, the sensitivity of the N170 to stimulus integrity (Bentin and Golland, 2002; Bentin et al., 2002) is consistent with hypotheses that the N170 is involved in higher level integration such as identification (Liu et al., 2002; Itier et al., 2006), a process not exclusive to faces.

Overall, our results stand in contrast to a large number of studies in the literature that have consistently reported face-selective N170 modulations. For instance, the N170 is increased in amplitude for inverted faces as compared to upright faces, independently of stimulus variability (Rossion et al., 2000; Boehm et al., 2011). However, it remains unexplained why the N170 should be increased in amplitude rather than reduced by inversion. It is intriguing that cropping, like inversion, increases N170 amplitude, perhaps because in both cases identification difficulty is increased. Visual expertise also has been repeatedly shown to modulate the N170 elicited by non-face–objects (Tanaka and Curran, 2001) or faces (Rossion et al., 2004) independent of cropping. However, effects of cropping on their own cannot explain all modulations found in theN170 range, just like inter-stimulus variance, symmetry, or other individual manipulations. Furthermore, since cropped faces arguably require being interpreted as faces, our results are not incompatible with conceptual priming effects such as those reported by Bentin et al. (2002) and Bentin and Golland (2002). Finally, we note that some studies have shown a lack of face-selectivity in the N170 range in congenitally prosopagnosic patients (Bentin et al., 2007a). However, Harris et al. (2005) have shown that the association between prosopagnosia and the absence of N170 face-selectivity is not straightforward, notwithstanding the fact that most experiments with prosopagnosic patients have systematically used cropped faces.

In sum, we establish that the N170 peak of visual event-related brain potentials is highly sensitive to stimulus integrity, i.e., it is increased in amplitude when stimuli are missing peripheral information, but fails to display category-selectivity with regard to the two contrast categories used here (cars and butterflies). Future studies will characterize the properties of the N170 that are potentially specific to face processing beyond the level of categorization (Eimer, 2000a; Itier and Taylor, 2002). More importantly, robust category-sensitivity regardless of low-level perceptual differences between conditions is consistently found in the P1 range, within 100 ms after picture presentation.

Chapter 5

N170 Modulation Is Expertise-Driven: Evidence from Word-Inversion Effects in Speakers of Different Languages

This chapter is accepted for publication in: CARLS proceedings: Future trends in the biology of language, Dering, B., Hoshino, N., & Thierry, G.

Abstract

Fifteen years of research have focused heavily on the N170 peak of event-related brain potentials to investigate electrophysiological correlates of face processing. The most established face-selective effect to date is the modulation of N170 amplitude by face inversion. Here, we show that a N170 modulation of the same magnitude as that recorded for faces is found in native readers of English for words presented upside-down. Moreover, this “word-inversion effect” is absent in non-native readers of English who have acquired a logographic language in childhood. These results demonstrate that the N170 inversion effect is not specific to faces and argue strongly in favor of an expertise-based account of the neural processes indexed by the N1.

Introduction

A substantial literature accumulated for more than forty years has widely established the human face as a special visual stimulus. Arguably, the strongest result supporting the specialness of faces is the so-called face-inversion effect (Yin, 1969). When faces are presented upside-down, they are much more difficult to recognize than other objects such as houses or furniture. Numerous studies in experimental psychology and neuropsychology, particularly with prosopagnosic patients who are specifically impaired for face recognition, have followed and demonstrated evidence for the existence of functional specialization in the human brain as regards face processing (Farah, 1996; Farah, Wilson, et al., 1995; Moscovitch, 1997).

The rationale for presenting faces in an abnormal orientation is to test the disruption of cognitive processes specific to upright faces such as holistic and configural processing (see Farah et al., 1995; Maurer, et al., 2002). When faces are inverted, holistic analysis is difficult, inducing a switch to a feature-based analysis by parts. Other authors, however, have extensively tested the hypothesis that the specialness of faces may be an emergent phenomenon due to the substantial expertise for faces developed by humans throughout development and learning. To demonstrate that expertise drives apparent face-selective effects, Gauthier & Tarr (1997) trained participants to recognize “greebles”, i.e., computer-generated novel objects differing from one another by visual attributes. These authors found greeble-inversion effects in individuals that became greeble experts whereas individuals new to greeble classification showed no inversion effects on greebles. Expertise effects also exist with structured object categories other than faces, such as dogs or birds (Diamond & Carey, 1986; Tanaka & Curran, 2001).

More recently, non-invasive brain imaging and electrophysiological techniques have been used to gain further insights into the functional specialization / expertise account of face recognition. For instance, using functional magnetic resonance imaging (fMRI), Kanwisher, et al. (1997) have shown that an area of the right midfusiform gyrus –also known as the Fusiform Face Area (FFA)– is selectively active when participants perceive faces as compared to other objects or scenes. Concurrently, in the field of event-related potentials (ERPs), Bentin and colleagues have established a selective modulation of the N1 component, termed the N170, which is larger in amplitude to the human face than any other object (Bentin, et al., 1996; Itier & Taylor, 2004a). Results from single-cell recordings in monkeys also have suggested the existence of neurons with face-selective patterns of response (Perrett, et al., 1984; Wang, Tanaka, & Tanifuji, 1996). However, there have been dissenting voices regarding the encapsulation of face processing with neuroimaging and electrophysiological techniques (Haxby, et al., 2001; Rossion, et al., 2000).

Focusing on the evidence from the ERP literature, a significant controversy has arisen regarding the face-selectivity of the N170. The N170 is a peak of activity characterized by a bilateral temporal negative deflection and a latency of around 170 ms typically larger in amplitude for faces than for objects from any other category. It has been repeatedly argued that the N170 reflects face-selective processes, possibly indicative of a structural encoding phase in processing faces (Bentin, et al., 1996; Eimer, 2000b; Eimer, et al., 2010; Itier & Taylor, 2004a) since structural encoding incorporates both early face categorization and identification processes (Bruce & Young, 1986). Recently, however, the face-selectivity of the N170 has been called into question (Boehm, Dering, & Thierry, 2011; Dering, Martin, Moro, Pegna, & Thierry, 2011; Dering, Martin, & Thierry, 2009; Rossion, et al., 2000; Schweinberger, et al., 2004; Thierry, et al., 2007a) and it has been suggested that category sensitivity to faces can be observed 70 ms earlier, in the range of the P1 component (Boehm, et al., 2011; Dering, et al., 2011; Dering, et al., 2009; Herrmann, Ehlis, Ellgring, et al., 2005; Thierry, et al., 2007a).

Moreover, Thierry et al. (2007a) were the first authors to describe conditions under which the N170 would not show face sensitivity, and warned that poor control of stimulus variance presented in different conditions might have confounded results and may make previous N170 observations inconclusive. These findings were attacked by a group of N170 experts (Bentin, et al., 2007) and despite a reply (Thierry, et al., 2007b), Rossion & Jacques (2008) published a review in an attempt to demonstrate that Thierry et al.'s (2007a) results were flawed and explain why the N170's largest amplitude to faces remains an established fact. In a recent paper by our group (Dering et al., 2011), we have pointed to further methodological problems with the interpretation of the N170 as a face-selective component of ERPs, going beyond the issue of stimulus variance raised by Thierry et al. (2007a) and extending into stimulus cropping.

One intriguing and yet unchallenged finding is that when faces are presented upside-down, the N170 is substantially amplified and delayed as compared to that elicited by upright faces and any other stimulus (Bentin, et al., 1996; Boehm, et al., 2006; Eimer, 2000a; Itier, et al., 2006; Itier & Taylor, 2002, 2004a), mimicking the behavioural face-inversion effect (Yin, 1969). This result has been put forward as strong evidence in favor of face-selectivity in the N170 range (e.g., Bentin et al. 2007). More specifically, the increase of N170 amplitude has been proposed to reflect the cost in switching from holistic to feature-based analysis. It is noteworthy that none of the effects tested or discussed by Thierry and colleagues (e.g., inter-stimulus variability, symmetry, stimulus cropping, effects of morphing) can account for the N170 inversion effect (as has been repeatedly stated, e.g., Bentin et al, 2007, Rossion and Jacques, 2008).

However, there exist another type of stimulus that is known to elicit strong N170 responses, often with a left-lateralized topography: Written words (Maurer, Brandeis, & McCandliss, 2005; McCandliss, Cohen, & Dehaene, 2003; Rossion, et al., 2003). Interestingly, the debate regarding functional specialization versus expertise has been going as strong in the field of word reading as it is in the field of face processing. Some authors (e.g., McCandliss et al., 2003) have proposed that a portion of the midfusiform gyrus –in the left hemisphere this time– is specialized for reading, the visual Word Form Area (WFA), while others (Price & Devlin, 2003) have strongly advocated distributed processing and an expertise-based account. More to the point, the left lateralized N170 elicited by words has been used to make claims regarding functional specialization for written words (see Mercure, Dick, Halit, Kaufman, & Johnson, 2008), but, oddly enough, word inversion has never been directly tested.

Written words are a class of stimuli made of parts that can be processed both holistically (global reading) and by parts (by decoding morphemes, syllables, or letters), requires learning, and ultimately results in high-levels of expertise in the fluent reader. Surprisingly, to our knowledge, the effect of presenting words upside down with ERPs has never been tested. Here, we used a simple 2 x 2 within- x 2 between-subject experimental design, involving faces and words presented in normal or inverted orientations. To test whether inversion effects are driven by expertise, we tested two groups of participants with very different experience of alphabetic stimuli: Native speakers of English who learnt to read English from the age of five, and native speakers of Chinese or Japanese, who were exposed to alphabetic stimuli only much later in life (>12 years).

We hypothesized that if words and faces recruit similar functional mechanisms, an increase in N170 mean amplitude should be observed for both English words and faces. Furthermore, if this effect is expertise-based, participants with lower expertise of alphabetic stimuli (late learner of written English) should show a reduced inversion-effect for words in the N170 range, whereas the inversion effect for faces should be of similar amplitude in both participant groups.

Methods

Participants

Twenty-five native speakers of English (hereafter British participants, 17 females, 3 left handed, mean age = 23.7 years, SD = 5.3) and 25 native speakers of Chinese or Japanese (hereafter Asian participants, 11 females, 0 left handed, mean age = 26.6 years, SD = 4.9) gave written informed consent to participate in the experiment that was approved by the ethics committee of Bangor University. All British participants received education in a British institution and learnt to read from the age of ~5. Asian participants were exposed to written alphabetic stimuli after the age of 11 (mean reading proficiency on a 10 point Likert scale = 6.4; mean writing proficiency = 5.7; mean speaking proficiency = 5.8; mean comprehension proficiency = 5.8; mean length of stay in the UK = 3.7 years; self reported time spent using English = 35.3%). All participants had self-reported normal or corrected-to-normal vision and no known neurological diseases.

Stimuli

Sixty images of full front neutral expression faces were obtained from the Productive Aging lab's face database (Minear & Park, 2004). They were selected such that the viewpoint was controlled and they had similar levels of symmetry. They were converted to grayscale and transposed on a medium gray background (RGB values; Fig. 1). Faces produced a maximum of 5.5 horizontal and 7.6 vertical degrees of visual angle. Sixty English words were selected from the MRC psycholinguistic database (http://www.psy.uwa.edu.au/mrcdatabase/uwa_mrc.htm) such that they had a familiarity rating of over 400 and were 4 letters in length to minimize horizontal eye-movement artifacts in the EEG signal. Words were displayed in Arial font, 48 point, producing 3.47 horizontal and 1 vertical degrees of visual angle.

Procedure

Stimuli were presented at the centre of a 19” CRT computer monitor (display refresh rate 75Hz) in a normal orientation or inverted, resulting in 240 experimental trials interspersed with other filler stimuli embedded for the purpose of another experiment and fully randomized between participants. All stimuli were presented for 200 ms, with an inter-stimulus interval of 1200 ms, and participants were asked to categorize each stimulus as word or face (irrespective of orientation) by pressing keys on a stimulus response box. Response sides were counterbalanced between participants.

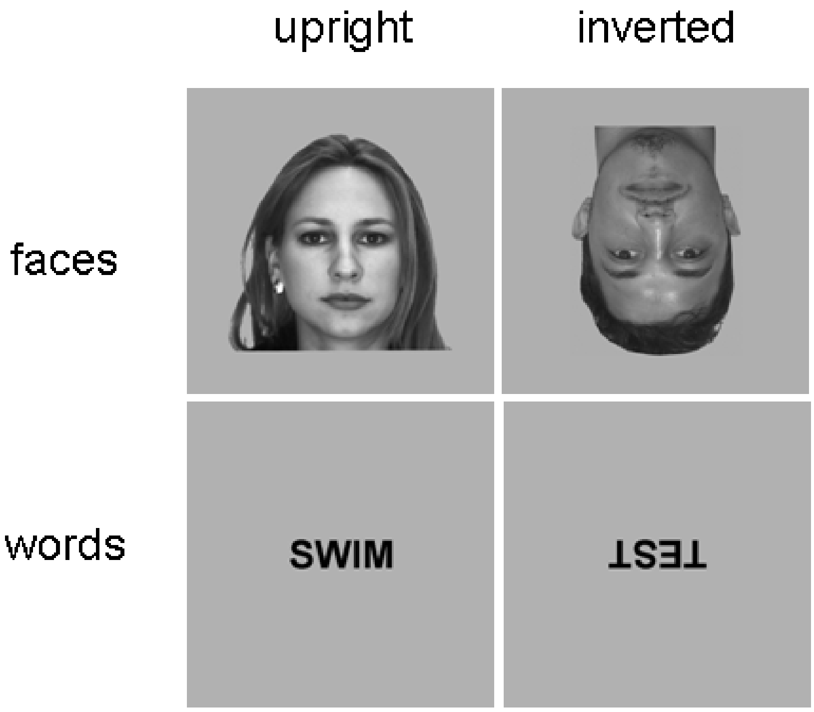

Figure 1 | **Examples of stimuli used in the experiment. Faces and words were presented randomly amid a stream with filler items.**

Event-related potentials

Electroencephalographic (EEG) signals were recorded from 64 Ag/AgCl electrodes distributed across the scalp according to the extended 10-20 system and sampled at 1 KHz. Cz was used as the reference during recording. Impedances were kept below 5 KΩ. EEG signals were filtered on-line between 0.01 and 200 Hz and off-line with a low-pass zero phase shift digital filter set to 30 Hz (48 db/octave slope). Eye-blink artifacts were mathematically corrected in each participant individually using a blink artifact modeling procedure (Gratton, Coles, & Donchin, 1983), and other non-blink artifacts exceeding ±75 µV were automatically discarded. Continuous EEG recordings were cut into epochs ranging from -200 ms to 800 ms after stimulus onset and averaged into ERPs for each individual in each of the 4 experimental conditions. Grand-averages were calculated after re-referencing individual ERPs to the common average reference. Individual mean amplitude time-windows were defined around the peak latency of the P1 and N1 for each condition. This method avoids the biasing of mean amplitude analyses by potential latency effects. P1 was identified as a positive peak between 80 – 120 ms and analyzed at sites P7, P8, PO7, PO8, PO9 and PO10. N170 peaked between 120 – 200 ms at electrode sites P7, P8, PO7, PO8, PO9 and PO10. Mean amplitude analyses for the P1 were conducted 20 ms around the peak of maximal activity for each condition (British participants peak latencies: upright faces 95 ms; inverted faces 99 ms; upright English words 90 ms; inverted English words 91 ms; Asian participants peak latencies: upright faces 89 ms; inverted faces 89 ms; upright English words 96 ms; inverted English words 93 ms). Mean amplitude analyses for the N170 were conducted 40 ms around the peak for each condition of the experiment (British participants peak latencies: upright faces 142 ms; inverted faces 150 ms; upright English words 143 ms; inverted English words 155 ms; Asian participants peak latencies: upright faces 140 ms; inverted faces 148 ms; upright English words 148 ms; inverted English words 149 ms). The data was subjected to repeated measures ANOVAs with 3 factors – category (face/car), inversion (upright/inverted) and electrode (6 levels). A Greenhouse-Geisser correction was used where applicable. To demonstrate the magnitude of effects, partial Eta squared (ήp2) is reported.

Results

Behavioural Results

Overall mean reaction time was 445 ms ±90 and overall mean accuracy was 90 ±12%. A repeated measures analysis of variance (ANOVA) showed that faces were overall categorized faster than written words of English [F(1,48)=11.41, p<.05, ήp2 = 0.192]. There was no other main effect or interaction on reaction times. Asian participants were more accurate than British participants in distinguishing categories [F(1,48)=7.18, p<.05, ήp2 = 0.13]. Inversion also affected accuracy [F(1,48)=5.44, p<.05, ήp2 = 0.102], such that Asian participants were more accurate with inverted than upright stimuli. There was no other effect.

ERP Results

ERPs displayed a typical P1-N1-P2 complex in all conditions (Fig. 2).

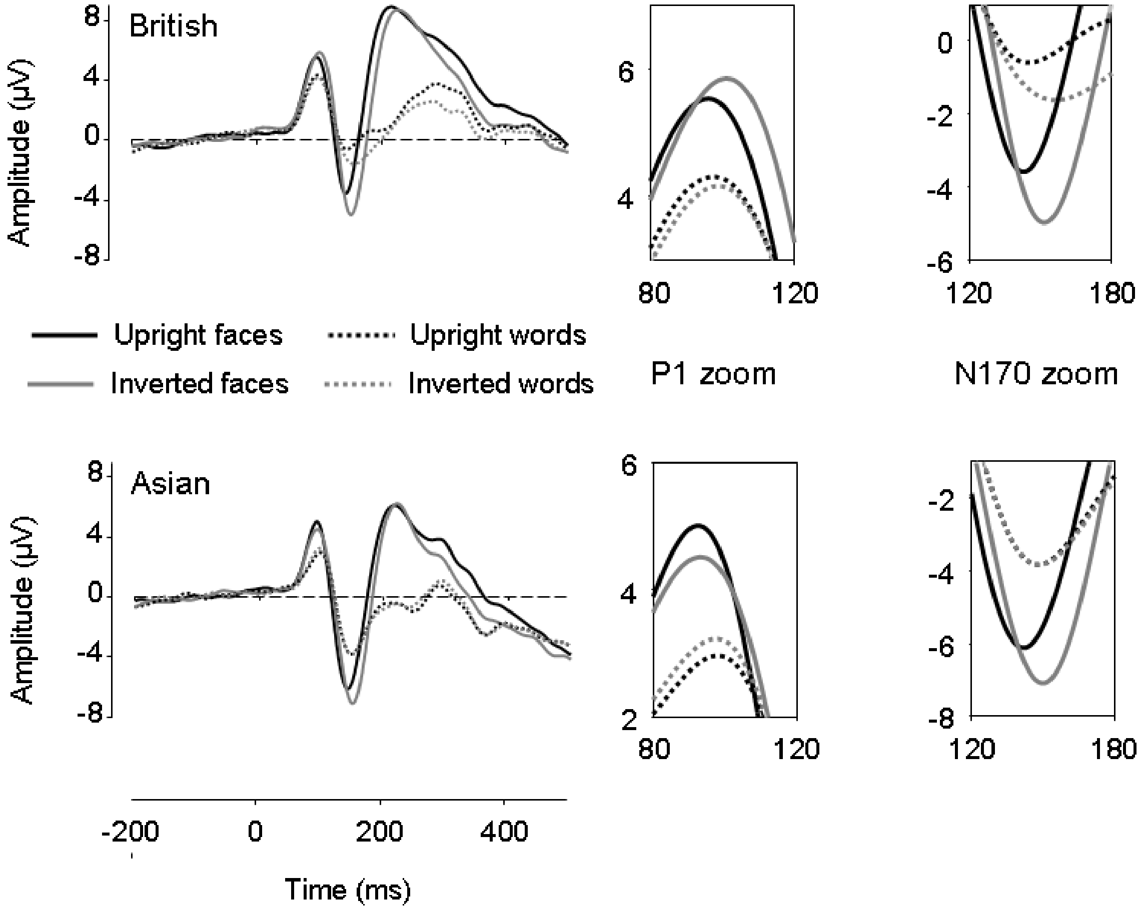

Figure 2 | **Grand averaged event-related brain potentials recorded in the four conditions of the experiment in British and Asian participants. Waveforms depict a linear derivation of the electrodes used in the statistical analysis for the P1 and N170, respectively**

P1

P1 mean amplitude was significantly different between categories [F(1,48)=37.4, p<.05, ήp2=0.438], with faces eliciting larger amplitudes than English words in both British and Asian participants. There was no effect of inversion on P1 mean amplitude [F(1,48)=.096, p>.05]. A three-way interaction between category, inversion and group showed that the P1 amplitude was modulated in opposing directions between groups [F(1,48)=4.36, p=.042, ήp2 = 0.083]: In the absence of a word inversion effect in the P1 range, British participants displayed larger amplitudes for inverted than upright faces whereas Asian participants showed the reverse pattern (larger amplitudes for upright than inverted faces).

P1 peak latencies were significantly shorter for faces than words [F(1,48)=105.186, p<.05, ήp2 = 0.687]. There was no main effect of inversion on P1 latencies [F(1,48)=1.497, p>.05], but there was an effect of group [F(1,48)=7.13, p<.05, ήp2 = 0.129], such that Asian participants had shorter P1 latencies than British participants. In addition, the group effect interacted with inversion [F(1,48)=10.63, p<.05, ήp2 = 0.181], such that P1 latency was delayed for inverted faces as compared to upright faces in British participants [F(1,26)=8.96, p<.05, ήp2 = 0.256] whereas no such delay was found in Asian participants.

N170

All three main effects were significant in the N170 range: We found a main effect of category on N170 mean amplitude [F(1,48)=29.398, p<.05, ήp2 = 0.38], such that faces produced overall larger N170 amplitudes than words. There was also a main effect of inversion showing that inverted stimuli elicited larger N170 than upright stimuli [F(1,48)=46.77, p<.05, ήp2 = 0.494]. There was also a main effect of group, such that Asian participants displayed larger N1 amplitudes than British participants [F(1,48)=4.739, p<.05, ήp2 = 0.09].

We found an inversion by group interaction indicating that the inversion effect on N170 was stronger in British participants than Asian participants [F(1,48)=8.811, p<.05, ήp2 = 0.155]. Furthermore and critically, a category by group interaction was found [F(1,48)=9.12, p<.05, ήp2 = 0.16], indicating differences between object categories and groups. Further exploration of this interaction revealed that written word inversion did not elicit any amplitude differences in Asian participants [F(1,22)=0.606, p>.05] whereas the inversion effect for words had the same magnitude as the inversion effect for faces in British participants [F(1,26)=15.23, p<.05, ήp2 = 0.369].

N170 peak latencies were overall shorter for upright as compared to inverted stimuli [F(1,48)=81.16, p<.05, ήp2 = 0.628]. There was also a main effect of category on N170 latencies [F(1,48)=7.96, p<.05, ήp2 = 0.142], such that written words elicited longer latencies than faces. Group interacted with inversion, such that Asian participants had shorter differences in latencies between upright and inverted stimuli than monolinguals [F(1,48)=10.12, p<.05, ήp2 = 0.174]. A three-way interaction between category, inversion and group also indicated that the shorter latencies for Asian participants were driven by the absence of latency lag between upright and inverted written words [F(1,48)=8.35, p<.05, ήp2 = 0.148].

Discussion

The aim of this study was to compare the effect of stimulus inversion on the processing of faces and written words of English using event-related potentials. Given the parallel functional considerations regarding face and word processing described in the introduction, we hypothesized that processing in the N170 range would be equally disrupted by stimulus inversion for faces and written words, resulting in an increase in amplitude and delayed peak latencies for inverted as compared to upright stimuli for both faces and words. As expected, British participants displayed an inversion effect of comparable magnitude in the N170 range for English words and faces. In addition, Asian participants who had moderate experience with alphabetic stimuli displayed no inversion effect for words and an inversion effect for faces comparable to that found in British participants.

Previous studies have consistently shown that faces and words have differently lateralised N170 peaks, with the N170 elicited by faces often reported as prominent over the right hemi-scalp (Bentin, et al., 1996; Boehm, et al., 2011) while the N170 elicited by written words is larger over left parietal areas (Maurer, Rossion, & McCandliss, 2008). This has led to claims that the N170 is category specific for faces and words, however, studies directly comparing these two classes of stimuli support an expertise-based account of face and word processing (Mercure et al., 2008). We found larger ERP amplitude for faces than written words of English in both the P1 and the N170 ranges. There are many parameters that may explain why such differences were found. Written words and faces differed along many dimensions such as spatial frequency, global and local contrast, luminance, size, overall shape, etc. and there is no reason why words and faces should have elicited comparable amplitudes at any point in time. ERPs elicited by stimuli of systematically different sizes for instance, are trivially increased in amplitude for the stimuli with the largest surface area related to the surface area of the retina being stimulated (Busch, et al., 2004; De Cesarei & Codispoti, 2006). Differences in luminance have also been shown to result in amplitude differences in the P1 range (e.g., Thierry, et al., 2009) as well as systematic differences in contrast (Eimer, 2011), and spatial frequency (Goffaux, et al., 2003; Goffaux & Rossion, 2006; Rousselet, et al., 2005). However, the overall difference in amplitude between the N170s elicited by faces and written words are of no importance in the context of the present study, because the critical hypotheses concerned interactions between stimulus category, inversion and participant group rather than main effects of object category.

Focussing on the inversion effects in the N170 range, it has been suggested that N170 amplitude increase indexes a switch from holistic to feature-based processing of faces (Jacques & Rossion, 2010). Since we found inversion effects for written words of comparable amplitude as that observed for faces in British participants, the same inference may be applied to the case of written words: Participants faced with words presented upside-down might be unable to resort to holistic processing as has been shown for normally oriented words (Grainger & Whitney, 2004; Pelli & Tillman, 2007) and would therefore switch to an analysis of smaller chunks to enable recognition. We speculate that an effect of task may even be observed if the same experiment was conducted in which a familiarity judgment on the specific faces and words was required, since determining the identity of a face or word is likely to increase processing demands significantly for inverted stimuli.

Another key finding of the present study is the absence of an inversion effect for written words of English in Asian participants who only learnt to read alphabetic stimuli after the age of 12 and therefore had accrued much less expertise with such stimuli than British participants. The expertise-based account of face processing contends that faces are a special class of stimulus because of people's extensive perceptual experience with faces. Other classes of objects have been shown to elicit similar inversion effects when the individuals tested were experts with these objects. For instance, Tanaka & Curran (2001) found expertise-driven N170 amplitude increases for dog and bird stimuli in dog and bird experts, respectively. Our results therefore suggest that the same functional principles can be applied to the case of written words. Indeed, a recent paper by Wang, Kuo, & Cheng (2011), has demonstrated that the inversion effect is found for Chinese characters in expert readers of Chinese. Since the N170 inversion effect is abolished for alphabetic stimuli in non-expert readers of English, the inversion effect is likely to index expertise-based processing of visual stimuli and there is no reason why such interpretation should not be generalised to the case of faces (Gauthier, Behrmann, et al., 1999; Gauthier & Tarr, 1997).

There remains one result difficult to interpret in our data: The reversal of P1 modulation between British and Asian participant groups. Whilst the larger P1 amplitude to upside-down relative to upright faces has been observed a number of times before (Itier and Taylor, 2004; Marzi & Viggiano, 2011), the observation of a reversal of this effect in Asian participants is a new, unpredicted finding. One interesting observation is that all the faces presented in our experiment were Caucasian faces, which Asian participants have arguably less experience of than Asian faces. It may be the case that the reversal of P1 modulation observed in the Asian participant group may be due to this differential experience with the specific stimuli used here, consistent with modulations by ethnic origin reported in the N170 range by others (Vizioli, et al., 2010). Further studies manipulating the ethnic origin of face stimuli are needed to address this particular question. We note however, that inversion effects in the P1 range were only found for faces and not words in both participant groups, a result consistent with previous conclusions from our group that P1 may be a better category-sensitive index than the N170 (Thierry et al., 2007a, Thierry et al., 2007b, Dering et al., 2009, Dering et al., 2011).

Conclusion

The increase in amplitude and peak delay of the N170 peak by stimulus inversion –commonly regarded as a face-selective effect (e.g., Bentin et al. 2007)– is also found for written words of English. The inversion effect, as indexed by the N170, is therefore not specific to faces. This conclusion gives further support to the view that N170 is not face-selective, whereas P1 continues to display a face-sensitive pattern of response (Herrman et al., 2005; Thierry et al., 2007a, Thierry et al., 2007b, Dering et al., 2009, Dering et al., 2011). In addition, Asian participants who only learnt to read in an alphabetic script after the age of 12 showed no N170 inversion effect for written words, although N170 modulation by inversion for faces was similar to that observed in British participants. We contend that this finding argues strongly in favour of an expertise-based account of the inversion effect, and supports the view that human faces are special because we are experts at processing them, rather than gifted with a functionally encapsulated and specialised device to process faces from birth.

Chapter 6

General discussion and concluding remarks

General Discussion

Summary of the overall results

The aim of this thesis was to test whether event-related potentials elicited by faces and other objects can index categorisation, i.e., the earliest stage of differentiation between classes of objects in visual processing. (1) In Chapter 2, we replicated Thierry et al. (2007a) and reported an N170 significantly larger to cars than faces over a set of electrodes commonly regarded as the optimal set for N170 analysis (Rossion and Jacques, 2008). Moreover, the task manipulation in (Dering, et al., 2009) failed to affect N170 amplitude. (2) Chapter 3 dealt with possible interactions between ISPV and inversion effects in the N170 range. While no inversion effect was observed for cars irrespective of ISPV, inverted increased N170 amplitude as expected, irrespective of ISPV in comparison to upright faces. (3) In chapter 4, cropping faces and cars elicits an artificial increase in N170 amplitudes, which may have led to the misinterpretation of face-selectivity in previous studies using cropped faces. In addition, the P1 was not significantly affected by cropping. (4) N170 amplitude was insensitive to cross-category morphing of faces and cars, whilst P1 displayed a face sensitive pattern of modulation. (5) Finally, in chapter 5 we showed that N170 inversion effects are of similar amplitude for faces and complex visual stimuli that require extensive expertise: written English words. Furthermore, participants with lower expertise of alphabetic stimuli (Asian participants) failed to show an inversion effect for words. Overall, and critically, in all six reported experiments, the P1 component, with a peak latency 70 ms earlier than the N170, was systematically larger in amplitude to faces than other objects.

On electrode selection

In a recent discussion of Thierry et al. (2007a), Rossion & Jacques (2008) argued that the N170 component in Thierry et al. (2007a, b) was measured at incorrect electrode sites. In addition to Thierry et al. (2007b)'s response to Bentin et al. (2007) showing that any choice of electrode would bring up the same results, Dering et al. (2009) found greater N170 amplitude for cars than faces, when analysed at the electrode sites recommended by Rossion & Jacques (2008), replicating the critical findings reported by Thierry et al. (2007a, b; see Fig. 1).

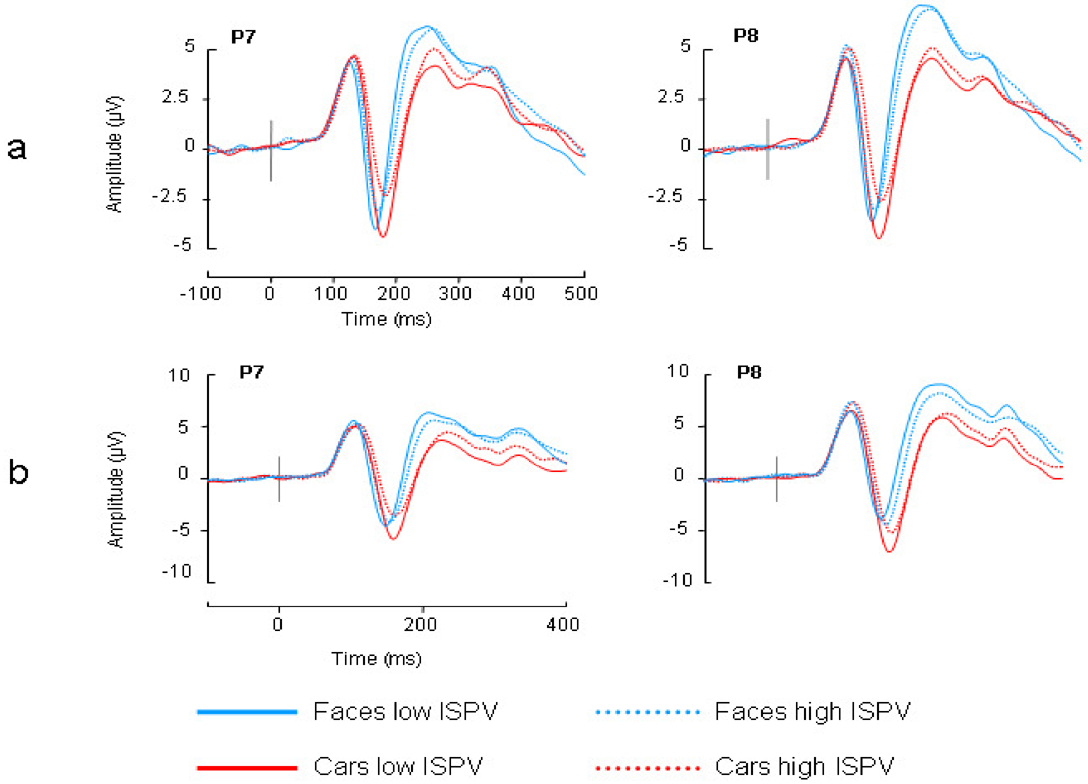

Figure 1 | **Here the ERP results of (a) Thierry et al. (2007a, b) and (b) those of Dering et al. (2009) are presented to show the replication of ISPV effects at electrode P7 and P8, often considered optimal to observe face-sensitivity. Note that N170 amplitude was significantly greater for cars as compared to faces in Dering et al. (2009) over a set of 6 electrodes, whereas the same effect was found only at electrode P8 in Thierry et al. (2007a, b).**

Contrary to what Rossion & Jacques (2008) claimed, there is no consensus regarding the number and location of the electrodes where N170 amplitudes should be measured. Rossion and Jacques (2008) chose to ignore that in Thierry et al. (2007b), studying the amplitude of the N170 at P7, P8, PO9 and PO10 (as requested by the reviewers and the editors of Nature Neuroscience in response to Bentin et al., 2007) replicated the pattern of results originally reported in Thierry et al. (2007a). The absence of sensitivity to category in the N170 range cannot be attributed to a “wrong choice of electrodes” since any choice within the array commonly accepted confirmed the results.

Expanding upon this issue of electrode selection, Rousselet, Pernet, Caldara, & Schyns, (2011) on Dering, et al. (2011) commented: *“Effects might also be distributed across electrodes, as multivariate multi-electrode analyses can reveal.”* However, these authors neglect the fact that scalp topographies were analysed in both Boehm, et al. (2011) and in Dering et al. (2011). In the latter, functional microstate topographies were affected by category in the P1 range but by cropping in the N170 range rather than category. Furthermore, source localisation in Boehm et al. (2011) indicated distinct neural generators for the N170 elicited by faces and cars in both the left and right hemispheres.

In sum, our position is that topographical differences are not fundamental when clear predictions exist as to the location of maximal differences for any given experimental contrast between conditions. When it is the case, prediction from previous studies can be directly tested at electrodes where the effects have been reported multiple times. Indeed, recall that our claim is limited to testing the hypothesis that the N170 is larger in amplitude for faces as compared to any other object category at the electrodes where it is maximal (e.g., P7, P8, PO9, PO10). Analyses conducted over the entire array of electrodes, single trials, and single participants (as also advocated by Rousselet et al., 2011), are clearly in the domain of empirical research and cannot be directly compared to the predictive approach taken in the present work.

Inter-stimulus variance has been controlled before

Both Rossion and Jacques (2008) and Eimer (2011) made the point that inter-stimulus variance has been controlled in many experiments before Thierry et al. (2007a) and Dering et al. (2011). Eimer (2011), for instance, contends that: *"…in the vast majority of published ERP studies on face processing, investigators have taken great care to equate the size, location, contrast, spatial frequency, and viewpoint for images of face and non-face objects, in order to minimize any differences between object categories…"* Whilst Eimer (2011) does not cite evidence to substantiate this claim, Rossion & Jacques (2008) provide a number of references in support of this point, and further added that *"…many other studies for which there is no reason to* *assume that [ISPV] was higher for nonface objects than faces also found a larger N170 to faces."*

We reviewed one-by-one the references listed in support of the claim by Rossion & Jacques (2008). Of the studies cited, only *one*, controlled for position, size, and viewpoint of the stimuli: Rousselet et al. (2005). However, Botzel et al. (1995) and Rousselet et al. (2004) did not control for position; Bentin et al. (1996), Itier and Taylor (2004a), Itier et al. (2006), and Rousselet (2004) did not control for size; Bentin et al. (1996), Carmel and Bentin (2002), Botzel et al. (1995), Itier and Taylor (2004a), Itier et al. (2006), Rossion et al. (2000), Rossion et al. (2003), Philiastides & Sajda (2006), and Rousselet et al. (2004a) did not control for viewpoint. Halgren, Raij, Marinkovic, Jousmaki, & Hari, (2000) and many other studies including one from our group (Thierry et al., 2006) did not control for category homogeneity (the object category involved a mix a various household objects). In addition, none of the studies cited compared centred, full-front (i.e. symmetrical) faces with objects that were all equally symmetrical (this includes, for instance, Eimer, 1998, 2000a, 2000b; Rousselet et al., 2005; and Zion-Golumbic & Bentin, 2007). Other studies apparently controlling for these factors used cropped images of faces (e.g., Boutsen, Humphreys, Praamstra, & Warbrick, 2006), while Herrman et al. (2005) found a larger P1 for faces, as well as N170 differences. To our knowledge, the only study other than ours in which faces, cars, and butterflies were controlled for position, size, and viewpoint, and were symmetrical, is that by Schweinberger et al. (2004) who found no difference between stimulus categories in the N170 range. In other words, the claim put forward by Rossion and Jacques (2008) and Eimer (2011) that perceptual variance was virtually always controlled in most previous studies is unfounded and false.

Rapid adaptation: The solution to physical differences?

Eimer (2011) highlights rapid adaptation as the solution to establish face-sensitivity in the N170 range and study the manipulations affecting N170 amplitude. The principle of this methodology is to present an adaptor stimulus(essentially a prime) followed by a test stimulus, which is always identical, making every condition ultimately comparable. Studies using rapid adaptation designs (Eimer, Gosling, Nicholas, & Kiss, 2011; Eimer, et al., 2010) have found that N170 to the test stimulus is significantly reduced in amplitude when the preceding adaptor is a face, compared to when the adaptor is another object, e.g., house-derived control stimuli (e.g., symbolic house or round section of a façade). Eimer et al. (2010) conclude that the highly specific pattern of N170 adaptation elicited by face adaptors reflects the activation of face-selective neurons in the absence of low-level visual differences between experimental conditions. However, there remains the possibility that this design involves significant confounds given that, even though the target stimulus is identical in all conditions, the adaptor stimulus is different between conditions (by definition) and necessarily affects the context of test stimulus presentation, which can result in baseline drifts or carry over effects impacting the processing of the test stimulus (see Steinhauer & Drury, 2011; for the same consideration in the domain of phase structure violations in language processing).

In both studies by Eimer et al. (2010; 2011), the P1 component preceding N170 was entirely neglected in the analysis, even though this component in their papers seems to be modulated by adaptation. Unfortunately, none of the electrodes where P1 is traditionally maximal in amplitude in response to face stimuli are reported or statistically studied. Therefore, it is impossible to know whether P1 modulation were significant and related to the patterns of response reported in the present thesis. In addition, the N170 was analysed at only two electrode locations (P7 & P8) in Eimer et al. (2010) and only at P8 in Eimer et al. (2011) and one wonders the extent to which effects found on peak amplitudes at one or two electrode are sufficiently robust. Strikingly, to our knowledge, adaptation studies conducted so far have exclusively tested faces as test stimuli (Eimer et al. 2010, 2011 in particular). It is therefore questionable how claims about category-selectivity can be made when only faces have been tested as test stimuli. Indeed, there is no theoretical reason why non-face objects (such as cars, for instance) would not produce similar adaptation effects as faces.

P1 face-sensitivity and physical differences

In their paper of 2005, Herrmann et al. (2005a) noted that the P1 component is often neglected in experiments reporting N170 category selectivity (e.g., Bentin et al.,1996; Bentin & Golland, 2002; Carmel & Bentin, 2002; Eimer, 1998, 2000a, b; Eimer et al. 2010; Eimer et al., 2011; Galli, Feurra, & Viggiano, 2006; Itier et al., 2006; Rossion & Jacques, 2008). Some studies have even used peak-to-peak analyses to study the N170, thereby cancelling any effect occurring before the N170 range (e.g., Kuefner, et al., 2010; Sadeh, et al., 2008). However, a growing body of evidence now considers the P1 component, peaking 70 ms before the N170, as face-sensitive (Linenkaer-Hansen et al., 1998; Iter and Taylor, 2004; Herrmann et al., 2005a,b; Thierry et al., 2007a,b; Dering et al., 2009; Boehm et al., 2011; Dering et al., 2011).

Recent publications have questioned this view of face-sensitivity in the P1 range (Kuefner, et al., 2010; Rossion & Caharel, 2011) using phase-scrambled as compared to unaltered images of faces and cars. Phase scrambling distributes phase values of the image randomly, while keeping intact amplitude spectrum, thus preserving the original global low-level physical properties (see Fig. 2). In both the studies cited above, the P1 was larger in response to faces, but also for phase-scrambled faces compared to phase-scrambled cars, leading the authors to argue that apparent P1 sensitivity to faces is entirely attributable to physical differences between faces and cars. However, the use of phase scrambling is disputable because it fails to eliminate physical differences between images of objects from different categories entirely due to non-random phase redistribution (see Fig. 2).

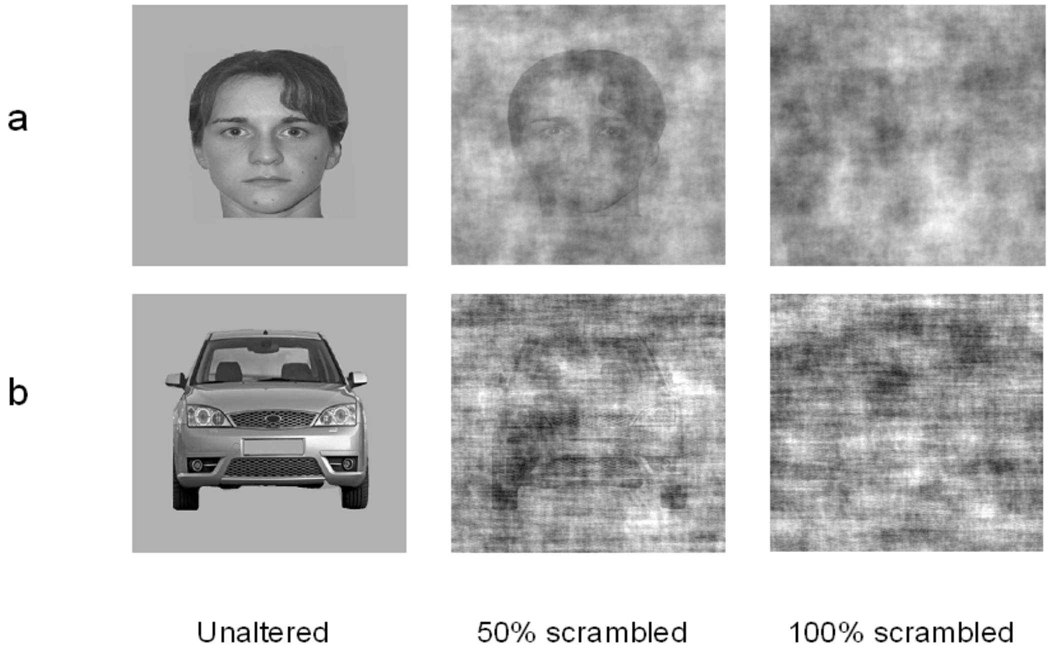

Figure 2 | **Illustration of the relationship between object category and phase-scrambled stimuli, showing that the scrambled stimuli are easily relatable to their original category. (a) Unaltered face, 50% phase scrambled face and 100% phase scrambled face; (b) Unaltered car, 50% phase scrambled car and 100% phase scrambled car. While phase scrambling preserves original global low-level visual properties, it disturbs object shape. Face and car can still be perceived in the 50% scrambled condition. In the 100% conditions, however, almost all shape information has been lost, yet local differences in contrast remain – high spatial frequency information is least affected by phase-scrambling compared to lower frequencies.**

Moreover, in both these studies, the authors chose to use colour images of faces and cars, which resulted in inherent colour differences in the phase-scrambled stimuli (see fig. 2 of Keufner et al. 2010; fig. 1 of Rossion & Caharel, 2011; and fig. 2 here). Finally, the authors did not control for low-level physical properties of the face and car images before scrambling. These three uncontrolled parameters lead to the strong possibility that participants quickly learnt –explicitly or implicitly– to categorise the scrambled stimuli as belonging to the same category as that of the non-scrambled object, thereby showing differences in the P1 range that are still category-selective rather than purely driven by low-level physical differences. In other words, P1 differences shown by Keufner et al. (2010) and Rossion & Caharel (2011) may have nothing to do with low-level physical differences between categories despite the strong claims put forward by these authors. Critically, the argument put forward by these authors is wholly inconsistent with the systematic observation of insensitivity of the P1 to inter-stimulus perceptual variance in Thierry et al. (2007a), Dering et al. (2009), Boehm et al. (2011), or to stimulus cropping (which arguably has a major effect of low-level physical properties, Dering et al. 2011).

Face processing: Modular or expertise driven?

In Chapter 2, we briefly reported a task effect on P1 amplitude independent of category-selective effects. Indeed, when participants were asked to categorise each presented stimulus as a face or car, P1 amplitude was boosted in comparison to the one-back task context. Assuming that faces are processed by an encapsulated module, which would have relative functional independence and be cognitively impenetrable, top-down modulation of its function driven by task requirement should not be observed (Fodor, 1983; Pylyshyn, 1999). Task effects have already been established in the N170 range (Crist, et al., 2007; Galli, et al., 2006; Righart, Burra, & Vuilleumier, 2011). The first observation of task effects in the P1 range could be considered evidence in support of global functional integration of face processing mechanisms rather than local modular encapsulation. Beyond the issue of modularity, results reported in Chapter 5 strongly suggest that expertise with visually complex stimuli based on early exposure leads to inversion effects with a stimulus class that has nothing to do with faces: written words of English. Therefore, the fundamental result provided here is not just that the N170 inversion effect can be found for other stimuli than faces, but rather that expertise is the source of the inversion effect (Rossion, et al., 2000). This suggests that expertise effects may be inherently linked with complex visual stimuli learnt over a long period of life, starting at an early age, and possibly concern socially relevant and meaningful stimuli rather than artificially learnt categories of novel meaningless objects.

Limitations and avenues for further study

Inter-stimulus perceptual variability modulates the amplitude of the N170 component, yet this effect varies between studies reported here and previously (Thierry et al., 2007a). Thierry et al. (2007a) found no significant differences in N170 amplitude for low variance faces and cars. Boehm et al. (2011) replicated this result on electrode sites P7 & P8, but not PO9 & PO10 and Dering et al. (2009) found greater N170 amplitude for cars than faces. These differences between experiments may be explained by group differences between the studies, and highlight the instability of the N170 response pattern. Rousselet et al. (2011) criticised Thierry et al. (2007a) and Dering et al. (2011) for using an experimental approach that is not sensitive enough because direct categorical contrasts always involve physical differences and claims should be made on the basis of interactions (Schyns et al, 2007). However, these limitations apply to all previous studies using direct category contrasts to study the N170, which represents essentially most of the literature. We note once again, however, that the P1 remained category-sensitive in simple categorical contrasts independently of any other factor manipulation throughout the work presented here and also in Thierry et al. (2007a,b).

Another issue is that ISPV is a factor difficult to quantify. Thierry et al. (2007a) provided pixel-wise correlations as a measure of physical variance of faces, cars, and butterflies, but they noted that this information may have no psychological relevance (cf. legend of Figure 2). It remains to be determined if perceptual variance and physical variance can be entirely dissociated (cf. Fig. 4 of the inroduction). Furthermore, controls for low-level physical properties of stimuli are only ever reported for the entire image presented in each trial, as in Dering et al. (2011). The image of a face with a low contrast, for instance, may be compensated by a darker background so that global contrast is matched to that of a face with higher local contrast and a lighter background. In figure 3 below, the face on the left has the exact same contrast as that on the right, even though perceptually, they are extremely different.

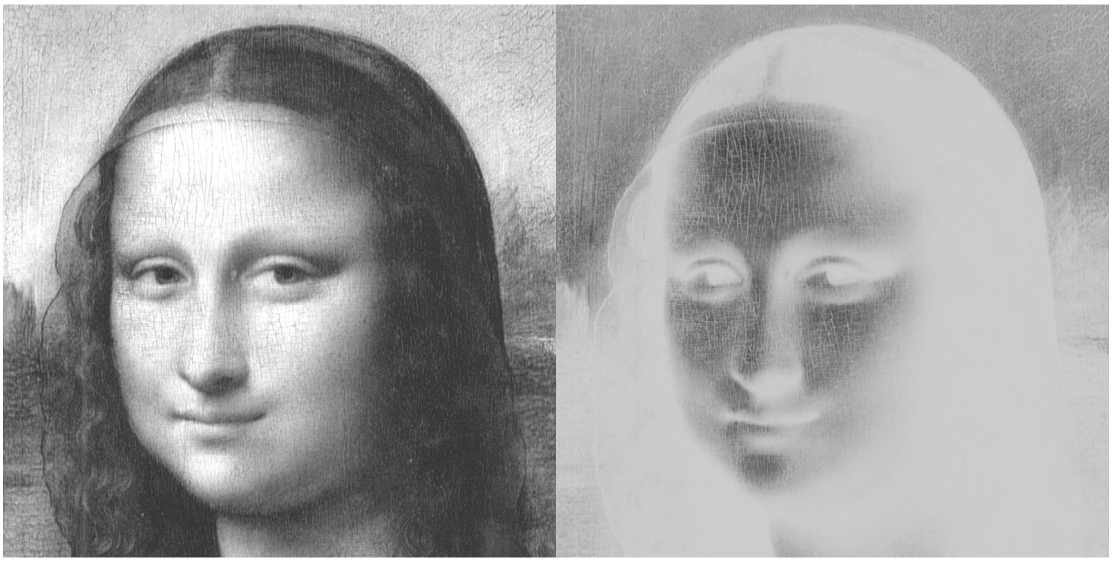

Figure 3 | **By inverting a picture of Mona Lisa, relative contrast is preserved exactly but perception is drastically altered.**

One suggestion put forward by Thierry et al. (2007a) in their original paper on ISPV is to use the exact same physical stimulus in each trial but manipulate where attention is focussed (see Fig. 3 of their paper in which cars and faces were overlaid in two different colours).

Finally, in all the experiments reported here, even though we never used recognition tasks, we found amplitude modulations of the N170, which may have been driven by recognition processes. In Dering et al. (2011), for instance, the increase in N170 amplitude elicited by cropping is possibly due to difficulty with face recognition but this hypothesis cannot be tested until the same experiment is conducted using a recognition task. It could be argued that this increase is observed because the identification of cropped images is more difficult than that of unaltered images, or because cropping affects object saliency due to unnatural appearance.

Concluding remarks

1. N170 face-selectivity is disputable since at least two categories of objects other than faces elicit peaks of indistinguishable magnitude when a number of low-level perceptual properties have been controlled between conditions (size, position, viewpoint, orientation, symmetry, luminance, and contrast, Dering et al. 2011).

2. Low-level physical differences between stimuli presented in an experiment must not be confused with perceptual differences. ISPV, for instance, confounds physical and perceptual properties and cannot be simply reduced to a low-level physical property of stimuli.

3. ISPV does not account for N170 modulations (and Thierry et al. 2007a never claimed it does) but ISPV definitely affects N170 amplitude in a significant fashion and more so for faces than other object categories.

4. The use of cropped faces in comparison to unaltered object categories should not be used to make claims regarding category selectivity, since cropped faces dramatically affect ERP peak amplitudes, whilst the stimulus presented is still clearly a face.

5. Even though it has seldom been observed in previous studies, the P1 shows category-sensitivity 70 ms before the N170 range and is unaffected by manipulations of ISPV and cropping.

6. Morphing objects from different categories shows that chimerical objects generated by morphing do not have intermediate status but rather are categorised as objects from one or the other category with “noise” from the other category.
7. Visually complex stimuli requiring high levels of expertise, such as faces and written English words, elicit similar patterns of ERP modulation by inversion, suggesting that inversion effects are generic and not specific to faces.

Future studies of face processing with ERPs should aim to dissociate between stages of visual object processing by studying the modulation of P1, N170, and later components, and determine whether faces are indeed processed in a different way to other complex visual objects.

References

Allison, T., Ginter, H., McCarthy, G., Nobre, A. C., Puce, A., Luby, M., et al. (1994). Face recognition in human extrastriate cortex. *J Neurophysiol, 71*(2), 821-825.

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*(5), 415-430.

Anaki, D., & Bentin, S. (2009). Familiarity effects on categorization levels of faces and objects. *Cognition, 111*(1), 144-149.

Andrews, T. J., Davies-Thompson, J., Kingstone, A., & Young, A. W. (2010). Internal and external features of the face are represented holistically in face-selective regions of visual cortex. *J Neurosci, 30*(9), 3544-3552.

Behrmann, M., Avidan, G., Marotta, J. J., & Kimchi, R. (2005). Detailed exploration of face-related processing in congenital prosopagnosia: 1. Behavioral findings. *J Cogn Neurosci, 17*(7), 1130-1149.

Bennett, T. L. (1992). Cognitive effects of epilepsy and anticonvulsant medications. In T. L. Bennett (Ed.), *The Neuropsychology of Epilepsy* (pp. 73-95). New York: Plenum Press.

Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *J Cogn Neuroci*(8), 551-565.

Bentin, S., Deouell, L. Y., & Soroker, N. (1999). Selective visual streaming in face recognition: evidence from developmental prosopagnosia. *Neuroreport, 10*(4), 823-827.

Bentin, S., & Golland, Y. (2002). Meaningful processing of meaningless stimuli: the influence of perceptual experience on early visual processing of faces. *Cognition, 86*(1), B1-14.

Bentin, S., Taylor, M. J., Rousselet, G. A., Itier, R. J., Caldara, R., Schyns, P. G., et al. (2007). Controlling interstimulus perceptual variance does not abolish N170 face sensitivity. *Nat Neurosci, 10*(7), 801-802; author reply 802-803.

Biederman, I. (1987). Recognition-by-components: a theory of human image understanding. *Psychol Rev, 94*(2), 115-147.

Bilalic, M., Turella, L., Campitelli, G., Erb, M., & Grodd, W. (2011). Expertise modulates the neural basis of context dependent recognition of objects and their relations. *Hum Brain Mapp*.

Blau, V. C., Maurer, U., Tottenham, N., & McCandliss, B. D. (2007). The face-specific N170 component is modulated by emotional facial expression. *Behav Brain Funct, 3*, 7.

Boehm, S. G., Dering, B., & Thierry, G. (2011). Category-sensitivity in the N170 range: A question of topography and inversion, not one of amplitude. *Neuropsychologia, 49*(7), 2082-2089.

Boehm, S. G., Klostermann, E. C., & Paller, K. A. (2006). Neural correlates of perceptual contributions to nondeclarative memory for faces. *Neuroimage, 30*(3), 1021-1029.

Botzel, K., Schulze, S., & Stodieck, S. R. (1995). Scalp topography and analysis of intracranial sources of face-evoked potentials. *Exp Brain Res, 104*(1), 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*(1), 352-367.

Braeutigam, S., Bailey, A. J., & Swithenby, S. J. (2001). Task-dependent early latency (30-60 ms) visual processing of human faces and other objects. *Neuroreport, 12*(7), 1531-1536.

Brandeis, D., & Lehmann, D. (1986). Event-related potentials of the brain and cognitive processes: approaches and applications. *Neuropsychologia, 24*(1), 151-168.

Bruce, V., & Young, A. (1986). Understanding face recognition. *Br J Psychol, 77 (Pt 3)*, 305-327.

Bruyer, R., & Crispeels, G. (1992). Expertise in person recognition. *Bulletin of the psychonomic society, 30*, 501 - 504.

Bukach, C. M., Gauthier, I., & Tarr, M. J. (2006). Beyond faces and modularity: the power of an expertise framework. *Trends Cogn Sci, 10*(4), 159-166.

Busch, N. A., Debener, S., Kranczioch, C., Engel, A. K., & Herrmann, C. S. (2004). Size matters: effects of stimulus size, duration and eccentricity on the visual gamma-band response. *Clin Neurophysiol, 115*(8), 1810-1820.

Busey, T. A., & Vanderkolk, J. R. (2005). Behavioral and electrophysiological evidence for configural processing in fingerprint experts. *Vision Res, 45*(4), 431-448.

Campanella, S., Hanoteau, C., Depy, D., Rossion, B., Bruyer, R., Crommelinck, M., et al. (2000). Right N170 modulation in a face discrimination task: an account for categorical perception of familiar faces. *Psychophysiology, 37*(6), 796-806.

Carmel, D., & Bentin, S. (2002). Domain specificity versus expertise: factors influencing distinct processing of faces. *Cognition, 83*(1), 1-29.

Crist, R. E., Wu, C. T., Karp, C., & Woldorff, M. G. (2007). Face Processing is Gated by Visual Spatial Attention. *Front Hum Neurosci, 1*, 10.

Damasio, A. R., Damasio, H., & Van Hoesen, G. W. (1982). Prosopagnosia: anatomic basis and behavioral mechanisms. *Neurology, 32*(4), 331-341.

De Cesarei, A., & Codispoti, M. (2006). When does size not matter? Effects of stimulus size on affective modulation. *Psychophysiology, 43*(2), 207-215.

De Renzi, E. (2000). Disorders of visual recognition. *Semin Neurol, 20*(4), 479-485.

Dering, B., Martin, C. D., Moro, S., Pegna, A. J., & Thierry, G. (2011). Face-sensitive processes one hundred milliseconds after picture onset. *Front Hum Neurosci, 5*, 93.

Dering, B., Martin, C. D., & Thierry, G. (2009). Is the N170 peak of visual event-related brain potentials car-selective? *Neuroreport, 20*(10), 902-906.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: an effect of expertise. *J Exp Psychol Gen, 115*(2), 107-117.

Doniger, G. M., Foxe, J. J., Schroeder, C. E., Murray, M. M., Higgins, B. A., & Javitt, D. C. (2001). Visual perceptual learning in human object recognition areas: a repetition priming study using high-density electrical mapping. *Neuroimage, 13*(2), 305-313.

Dubal, S., Foucher, A., Jouvent, R., & Nadel, J. (2011). Human brain spots emotion in non humanoid robots. *Soc Cogn Affect Neurosci, 6*(1), 90-97.

Duchaine, B. C., & Nakayama, K. (2004). Developmental prosopagnosia and the Benton Facial Recognition Test. *Neurology, 62*(7), 1219-1220.

Ebersole, J. S. (1997). Defining epileptogenic foci: past, present, future. *J Clin Neurophysiol, 14*(6), 470-483.

Eimer, M. (2000a). Effects of face inversion on the structural encoding and recognition of faces. Evidence from event-related brain potentials. *Brain Res Cogn Brain Res, 10*(1-2), 145-158.

Eimer, M. (2000b). The face-specific N170 component reflects late stages in the structural encoding of faces. *Neuroreport, 11*(10), 2319-2324.

Eimer, M. (2011). The face-sensitivity of the N170 component. *Front Hum Neurosci, 5*.

Eimer, M., Gosling, A., Nicholas, S., & Kiss, M. (2011). The N170 component and its links to configural face processing: a rapid neural adaptation study. *Brain Res, 1376*, 76-87.

Eimer, M., Kiss, M., & Nicholas, S. (2010). Response profile of the face-sensitive N170 component: a rapid adaptation study. *Cereb Cortex, 20*(10), 2442-2452.

Eimer, M., & McCarthy, R. A. (1999). Prosopagnosia and structural encoding of faces: evidence from event-related potentials. *Neuroreport, 10*(2), 255-259.

Farah, M. J. (1996). Is face recognition 'special'? Evidence from neuropsychology. *Behav Brain Res, 76*(1-2), 181-189.

Farah, M. J., Levinson, K. L., & Klein, K. L. (1995). Face perception and within-category discrimination in prosopagnosia. *Neuropsychologia, 33*(6), 661-674.

Farah, M. J., Tanaka, J. W., & Drain, H. M. (1995). What causes the face inversion effect? *J Exp Psychol Hum Percept Perform, 21*(3), 628-634.

Farah, M. J., Wilson, K. D., Drain, H. M., & Tanaka, J. R. (1995). The inverted face inversion effect in prosopagnosia: evidence for mandatory, face-specific perceptual mechanisms. *Vision Res, 35*(14), 2089-2093.

Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is "special" about face perception? *Psychol Rev, 105*(3), 482-498.

Fodor, J. A. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.

Galli, G., Feurra, M., & Viggiano, M. P. (2006). "Did you see him in the newspaper?" Electrophysiological correlates of context and valence in face processing. *Brain Res, 1119*(1), 190-202.

Gauthier, I., Behrmann, M., & Tarr, M. J. (1999). Can face recognition really be dissociated from object recognition? *J Cogn Neurosci, 11*(4), 349-370.

Gauthier, I., & Curby, K. M. (2005). A perceptual traffic jam on highway N170. *Psychological Science, 14*, 30 – 33.

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nat Neurosci, 3*(2), 191-197.

Gauthier, I., & Tarr, M. J. (1997). Becoming a "Greeble" expert: exploring mechanisms for face recognition. *Vision Res, 37*(12), 1673-1682.

Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (1999). Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects. *Nat Neurosci, 2*(6), 568-573.

Gauthier, I., Williams, P., Tarr, M. J., & Tanaka, J. (1998). Training 'greeble' experts: a framework for studying expert object recognition processes. *Vision Res, 38*(15-16), 2401-2428.

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*(2), 65-76.

Gobbini, M. I., & Haxby, J. V. (2007). Neural systems for recognition of familiar faces. *Neuropsychologia, 45*(1), 32-41.

Goffaux, V., Gauthier, I., & Rossion, B. (2003). Spatial scale contribution to early visual differences between face and object processing. *Brain Res Cogn Brain Res, 16*(3), 416-424.

Goffaux, V., & Rossion, B. (2006). Faces are "spatial"--holistic face perception is supported by low spatial frequencies. *J Exp Psychol Hum Percept Perform, 32*(4), 1023-1039.

Grainger, J., & Whitney, C. (2004). Does the huamn mnid raed wrods as a wlohe? *Trends Cogn Sci, 8*(2), 58-59.

Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalogr Clin Neurophysiol, 55*(4), 468-484.

Gschwind, M., Pourtois, G., Schwartz, S., Van De Ville, D., & Vuilleumier, P. (2011). White-Matter Connectivity between Face-Responsive Regions in the Human Brain. *Cereb Cortex*.

Hadjikhani, N., Kveraga, K., Naik, P., & Ahlfors, S. P. (2009). Early (M170) activation of face-specific cortex by face-like objects. *Neuroreport, 20*(4), 403-407.

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*(1), 69-81.

Harris, A. M., Duchaine, B. C., & Nakayama, K. (2005). Normal and abnormal face selectivity of the M170 response in developmental prosopagnosics. *Neuropsychologia, 43*(14), 2125-2136.

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science, 293*(5539), 2425-2430.

Heisz, J. J., Watter, S., & Shedden, J. M. (2006). Progressive N170 habituation to unattended repeated faces. *Vision Res, 46*(1-2), 47-56.

Herrmann, M. J., Ehlis, A. C., Ellgring, H., & Fallgatter, A. J. (2005). Early stages (P100) of face perception in humans as measured with event-related potentials (ERPs). *J Neural Transm, 112*(8), 1073-1081.

Herrmann, M. J., Ehlis, A. C., Muehlberger, A., & Fallgatter, A. J. (2005). Source localization of early stages of face processing. *Brain Topogr, 18*(2), 77-85.

Herzmann, G., Kunina, O., Sommer, W., & Wilhelm, O. (2009). Individual differences in face cognition: brain-behavior relationships. *J Cogn Neurosci, 22*(3), 571-589.

Hopf, J. M., & Mangun, G. R. (2000). Shifting visual attention in space: an electrophysiological analysis using high spatial resolution mapping. *Clin Neurophysiol, 111*(7), 1241-1257.

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*(4), 1587-1595.

Itier, R. J., Latinus, M., & Taylor, M. J. (2006). Face, eye and object early processing: what is the face specificity? *Neuroimage, 29*(2), 667-676.

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*(2), 353-372.

Itier, R. J., & Taylor, M. J. (2004a). N170 or N1? Spatiotemporal differences between object and face processing using ERPs. *Cereb Cortex, 14*(2), 132-142.

Itier, R. J., & Taylor, M. J. (2004b). Source analysis of the N170 to faces and objects. *Neuroreport, 15*(8), 1261-1265.

Itier, R. J., Taylor, M. J., & Lobaugh, N. J. (2004). Spatiotemporal analysis of event-related potentials to upright, inverted, and contrast-reversed faces: effects on encoding and recognition. *Psychophysiology, 41*(4), 643-653.

Jacques, C., & Rossion, B. (2007). Electrophysiological evidence for temporal dissociation between spatial attention and sensory competition during human face processing. *Cereb Cortex, 17*(5), 1055-1065.

Jacques, C., & Rossion, B. (2010). Misaligning face halves increases and delays the N170 specifically for upright faces: implications for the nature of early face representations. *Brain Res, 1318*, 96-109.

Jeffreys, D. A. (1996). Evoked potential studies of face and object processing. *Visual Cognition, 3*(1), 1-38.

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*(16), 2035-2039.

Jemel, B., Schuller, A. M., & Goffaux, V. (2010). Characterizing the spatio-temporal dynamics of the neural events occurring prior to and up to overt recognition of famous faces. *J Cogn Neurosci, 22*(10), 2289-2305.

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*(11), 4302-4311.

Kanwisher, N., & Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. *Philos Trans R Soc Lond B Biol Sci, 361*(1476), 2109-2128.

Koenig, T., & Melie-Garcia, L. (2010). A method to determine the presence of averaged event-related fields using randomization tests. *Brain Topogr, 23*(3), 233-242.

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*(5), 742-753.

Krolak-Salmon, P., Henaff, M. A., Vighetto, A., Bertrand, O., & Mauguiere, F. (2004). Early amygdala reaction to fear spreading in occipital, temporal, and frontal cortex: a depth electrode ERP study in human. *Neuron, 42*(4), 665-676.

Kuefner, D., de Heering, A., Jacques, C., Palmero-Soler, E., & Rossion, B. (2010). Early Visually Evoked Electrophysiological Responses Over the Human Brain (P1, N170) Show Stable Patterns of Face-Sensitivity from 4 years to Adulthood. *Front Hum Neurosci, 3,* 67.

Lehmann, D., & Skrandies, W. (1984). Spatial analysis of evoked potentials in man--a review. *Prog Neurobiol, 23*(3), 227-250.

Leleu, A., Caharel, S., Carre, J., Montalan, B., Snoussi, M., Vom Hofe, A., et al. (2010). Perceptual interactions between visual processing of facial familiarity and emotional expression: an event-related potentials study during task-switching. *Neurosci Lett, 482*(2), 106-111.

Linkenkaer-Hansen, K., Palva, J. M., Sams, M., Hietanen, J. K., Aronen, H. J., & Ilmoniemi, R. J. (1998). Face-selective processing in human extrastriate cortex around 120 ms after stimulus onset revealed by magneto- and electroencephalography. *Neurosci Lett, 253*(3), 147-150.

Liu, J., Harris, A., & Kanwisher, N. (2002). Stages of processing in face perception: an MEG study. *Nat Neurosci, 5*(9), 910-916.

Liu, J., Higuchi, M., Marantz, A., & Kanwisher, N. (2000). The selectivity of the occipitotemporal M170 for faces. *Neuroreport, 11*(2), 337-341.

Liu J, Tian, J., Lee, K., & Li, J. (2008). A study on neural mechanism of face processing based on fMRI. *Prog Nat Sci*, 18:201-207.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco, CA: W.H. Freeman.

Marzi, T., & Viggiano, M. P. (2011). Temporal dynamics of face inversion at encoding and retrieval. *Clin Neurophysiol, 122*(7), 1360-1370.

Maurer, D., Grand, R. L., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends Cogn Sci, 6*(6), 255-260.

Maurer, U., Brandeis, D., & McCandliss, B. D. (2005). Fast, visual specialization for reading in English revealed by the topography of the N170 ERP response. *Behav Brain Funct, 1*, 13.

Maurer, U., Rossion, B., & McCandliss, B. D. (2008). Category specificity in early perception: face and word n170 responses differ in both lateralization and habituation properties. *Front Hum Neurosci, 2*, 18.

McCandliss, B. D., Cohen, L., & Dehaene, S. (2003). The visual word form area: expertise for reading in the fusiform gyrus. *Trends Cogn Sci, 7*(7), 293-299.

McCarthy, G., & Wood, C. C. (1985). Scalp distributions of event-related potentials: an ambiguity associated with analysis of variance models. *Electroencephalogr Clin Neurophysiol, 62*(3), 203-208.

Mercure, E., Dick, F., Halit, H., Kaufman, J., & Johnson, M. H. (2008). Differential lateralization for words and faces: category or psychophysics? *J Cogn Neurosci, 20*(11), 2070-2087.

Mercure, E., Dick, F., & Johnson, M. H. (2008). Featural and configural face processing differentially modulate ERP components. *Brain Res, 1239*, 162-170.

Michel, C. M., Seeck, M., & Landis, T. (1999). Spatiotemporal Dynamics of Human Cognition. *News Physiol Sci, 14*, 206-214.

Michel, C. M., Thut, G., Morand, S., Khateb, A., Pegna, A. J., Grave de Peralta, R., et al. (2001). Electric source imaging of human brain functions. *Brain Res Brain Res Rev, 36*(2-3), 108-118.

Milivojevic, B., Clapp, W. C., Johnson, B. W., & Corballis, M. C. (2003). Turn that frown upside down: ERP effects of thatcherization of misorientated faces. *Psychophysiology, 40*(6), 967-978.

Milner, A. D., & Goodale, M. A. (1995). *The visual brain in action*. Oxford: Oxford University Press.

Minear, M., & Park, D. C. (2004). A lifespan database of adult facial stimuli. *Behav Res Methods Instrum Comput, 36*(4), 630-633.

Minnebusch, D. A., Suchan, B., Ramon, M., & Daum, I. (2007). Event-related potentials reflect heterogeneity of developmental prosopagnosia. *Eur J Neurosci, 25*(7), 2234-2247.

Mohamed, T. N., Neumann, M. F., & Schweinberger, S. R. (2009). Perceptual load manipulation reveals sensitivity of the face-selective N170 to attention. *Neuroreport, 20*(8), 782-787.

Moscovitch, M., Winocur, G., & Behrmann, M. (1997). What Is Special about Face Recognition? Nineteen Experiments on a Person with Visual Object Agnosia and Dyslexia but Normal Face Recognition. *Journal of Cognitive Neuroscience, 9*(5), 555 - 604.

Mouchetant-Rostaing, Y., & Giard, M. H. (2003). Electrophysiological correlates of age and gender perception on human faces. *J Cogn Neurosci, 15*(6), 900-910.

Murray, M. M., Brunet, D., & Michel, C. M. (2008). Topographic ERP analyses: a step-by-step tutorial review. *Brain Topogr, 20*(4), 249-264.

Nakashima, T., Kaneko, K., Goto, Y., Abe, T., Mitsudo, T., Ogata, K., et al. (2008). Early ERP components differentially extract facial features: evidence for spatial frequency-and-contrast detectors. *Neurosci Res, 62*(4), 225-235.

Nasr, S., & Esteky, H. (2009). A study of N250 event-related brain potential during face and non-face detection tasks. *J Vis, 9*(5), 5 1-14.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia, 9*(1), 97-113.

Oram, M. W., & Perrett, D. I. (1992). Time course of neural responses discriminating different views of the face and head. *J Neurophysiol, 68*(1), 70-84.

Pascual-Marqui, R. D., Michel, C. M., & Lehmann, D. (1995). Segmentation of brain electrical activity into microstates: model estimation and validation. *IEEE Trans Biomed Eng, 42*(7), 658-665.

Pegna, A. J., Khateb, A., Michel, C. M., & Landis, T. (2004). Visual recognition of faces, objects, and words using degraded stimuli: where and when it occurs. *Hum Brain Mapp, 22*(4), 300-311.

Pegna, A. J., Khateb, A., Spinelli, L., Seeck, M., Landis, T., & Michel, C. M. (1997). Unraveling the cerebral dynamics of mental imagery. *Hum Brain Mapp, 5*(6), 410-421.

Pelli, D. G., & Tillman, K. A. (2007). Parts, wholes, and context in reading: a triple dissociation. *PLoS One, 2*(8), e680.

Perrett, D. I., Oram, M. W., & Ashbridge, E. (1998). Evidence accumulation in cell populations responsive to faces: an account of generalisation of recognition without mental transformations. *Cognition, 67*(1-2), 111-145.

Perrett, D. I., Smith, P. A., Potter, D. D., Mistlin, A. J., Head, A. S., Milner, A. D., et al. (1984). Neurones responsive to faces in the temporal cortex: studies of functional organization, sensitivity to identity and relation to perception. *Hum Neurobiol, 3*(4), 197-208.

Philiastides, M. G., & Sajda, P. (2006). Temporal characterization of the neural correlates of perceptual decision making in the human brain. *Cereb Cortex, 16*(4), 509-518.

Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R., Jr., et al. (2000). Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria. *Psychophysiology, 37*(2), 127-152.

Pitcher, D., Dilks, D. D., Saxe, R. R., Triantafyllou, C., & Kanwisher, N. (2011). Differential selectivity for dynamic versus static information in face-selective cortical regions. *Neuroimage, 56*(4), 2356-2363.

Pitcher, D., Walsh, V., Yovel, G., & Duchaine, B. (2007). TMS evidence for the involvement of the right occipital face area in early face processing. *Curr Biol, 17*(18), 1568-1573.

Price, C. J., & Devlin, J. T. (2003). The myth of the visual word form area. *Neuroimage, 19*(3), 473-481.

Pylyshyn, Z. (1999). Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behav Brain Sci, 22*(3), 341-365; discussion 366-423.

Righart, R., Burra, N., & Vuilleumier, P. (2011). Face perception in the mind's eye. *Brain Topogr, 24*(1), 9-18.

Righart, R., & de Gelder, B. (2007). Impaired face and body perception in developmental prosopagnosia. *Proc Natl Acad Sci U S A, 104*(43), 17234-17238.

Robbins, R., & McKone, E. (2007). No face-like processing for objects-of-expertise in three behavioural tasks. *Cognition, 103*(1), 34-79.

Rossion, B. (2008). Picture-plane inversion leads to qualitative changes of face perception. *Acta Psychol (Amst), 128*(2), 274-289.

Rossion, B., & Caharel, S. (2011). ERP evidence for the speed of face categorization in the human brain: Disentangling the contribution of low-level visual cues from face perception. *Vision Res, 51*(12), 1297-1311.

Rossion, B., & Curran, T. (2010). Visual expertise with pictures of cars correlates with RT magnitude of the car inversion effect. *Perception, 39*(2), 173-183.

Rossion, B., Curran, T., & Gauthier, I. (2002). A defense of the subordinate-level expertise account for the N170 component. *Cognition, 85*(2), 189-196.

Rossion, B., Delvenne, J. F., Debatisse, D., Goffaux, V., Bruyer, R., Crommelinck, M., et al. (1999). Spatio-temporal localization of the face inversion effect: an event-related potentials study. *Biol Psychol, 50*(3), 173-189.

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*(3), 250-257.

Rossion, B., Gauthier, I., Tarr, M. J., Despland, P., Bruyer, R., Linotte, S., et al. (2000). The N170 occipito-temporal component is delayed and enhanced to inverted faces but not to inverted objects: an electrophysiological account of face-specific processes in the human brain. *Neuroreport, 11*(1), 69-74.

Rossion, B., & Jacques, C. (2008). Does physical interstimulus variance account for early electrophysiological face sensitive responses in the human brain? Ten lessons on the N170. *Neuroimage, 39*(4), 1959-1979.

Rossion, B., Joyce, C. A., Cottrell, G. W., & Tarr, M. J. (2003). Early lateralization and orientation tuning for face, word, and object processing in the visual cortex. *Neuroimage, 20*(3), 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*(40), 14521-14526.

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*(10), 755-763.

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*(3), 843-862.

Rousselet, G. A., Mace, M. J., & Fabre-Thorpe, M. (2004). Spatiotemporal analyses of the N170 for human faces, animal faces and objects in natural scenes. *Neuroreport, 15*(17), 2607-2611.

Rousselet, G. A., Pernet, C. R., Caldara, R., & Schyns, P. G. (2011). Visual object categorization in the brain: What can we really learn from ERP peaks? *Front Hum Neurosci, 5*(156).

Sadeh, B., Zhdanov, A., Podlipsky, I., Hendler, T., & Yovel, G. (2008). The validity of the face-selective ERP N170 component during simultaneous recording with functional MRI. *Neuroimage, 42*(2), 778-786.

Sagiv, N., & Bentin, S. (2001). Structural encoding of human and schematic faces: holistic and part-based processes. *J Cogn Neurosci, 13*(7), 937-951.

Saumier, D., Arguin, M., & Lassonde, M. (2001). Prosopagnosia: a case study involving problems in processing configural information. *Brain Cogn, 46*(1-2), 255-259.

Schiltz, C., Sorger, B., Caldara, R., Ahmed, F., Mayer, E., Goebel, R., et al. (2006). Impaired face discrimination in acquired prosopagnosia is associated with abnormal response to individual faces in the right middle fusiform gyrus. *Cereb Cortex, 16*(4), 574-586.

Schupp, H. T., Stockburger, J., Schmalzle, R., Bublatzky, F., Weike, A. I., & Hamm, A. O. (2008). Visual noise effects on emotion perception: brain potentials and stimulus identification. *Neuroreport, 19*(2), 167-171.

Schweinberger, S. R., Huddy, V., & Burton, A. M. (2004). N250r: a face-selective brain response to stimulus repetitions. *Neuroreport, 15*(9), 1501-1505.

Schweinberger, S. R., Kaufmann, J. M., Moratti, S., Keil, A., & Burton, A. M. (2007). Brain responses to repetitions of human and animal faces, inverted faces, and objects: an MEG study. *Brain Res, 1184*, 226-233.

Seeck, M., Michel, C. M., Blanke, O., Thut, G., Landis, T., & Schomer, D. L. (2001). Intracranial Neurophysiological Correlates Related to the Processing of Faces. *Epilepsy Behav, 2*(6), 545-557.

Seeck, M., Michel, C. M., Mainwaring, N., Cosgrove, R., Blume, H., Ives, J., et al. (1997). Evidence for rapid face recognition from human scalp and intracranial electrodes. *Neuroreport, 8*(12), 2749-2754.

Steinhauer, K., & Drury, J. E. (2011). On the early left-anterior negativity (ELAN) in syntax studies. *Brain Lang*.

Stollhoff, R., Jost, J., Elze, T., & Kennerknecht, I. (2010). The early time course of compensatory face processing in congenital prosopagnosia. *PLoS One, 5*(7), e11482.

Tacikowski, P., Jednorog, K., Marchewka, A., & Nowicka, A. (2011). How multiple repetitions influence the processing of self-, famous and unknown names and faces: an ERP study. *Int J Psychophysiol, 79*(2), 219-230.

Tanaka, J., & Curran, T. (2001). A neural basis for expert object recognition. *Psychol Sci, 12*(1), 43-47.

Tanaka, J., Luu, P., Weisbrod, M., & Kiefer, M. (1999). Tracking the time course of object categorization using event-related potentials. *Neuroreport, 10*(4), 829-835.

Taylor, J. C., & Downing, P. E. (2011). Division of Labor between Lateral and Ventral Extrastriate Representations of Faces, Bodies, and Objects. *J Cogn Neurosci, 23*(12), 4122-4137.

Thierry, G., Athanasopoulos, P., Wiggett, A., Dering, B., & Kuipers, J. R. (2009). Unconscious effects of language-specific terminology on preattentive color perception. *Proc Natl Acad Sci U S A, 106*(11), 4567-4570.

Thierry, G., Martin, C. D., Downing, P., & Pegna, A. J. (2007a). Controlling for interstimulus perceptual variance abolishes N170 face selectivity. *Nat Neurosci, 10*(4), 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.

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*(2), 871-879.

Tsao, D. Y., Freiwald, W. A., Tootell, R. B., & Livingstone, M. S. (2006). A cortical region consisting entirely of face-selective cells. *Science, 311*(5761), 670-674.

Urbach, T. P., & Kutas, M. (2002). The intractability of scaling scalp distributions to infer neuroelectric sources. *Psychophysiology, 39*(6), 791-808.

Vizioli, L., Rousselet, G. A., & Caldara, R. (2010). Neural repetition suppression to identity is abolished by other-race faces. *Proc Natl Acad Sci U S A, 107*(46), 20081-20086.

Vuilleumier, P., & Pourtois, G. (2007). Distributed and interactive brain mechanisms during emotion face perception: evidence from functional neuroimaging. *Neuropsychologia, 45*(1), 174-194.

Wang, G., Tanaka, K., & Tanifuji, M. (1996). Optical imaging of functional organization in the monkey inferotemporal cortex. *Science, 272*(5268), 1665-1668.

Wang, M. Y., Kuo, B. C., & Cheng, S. K. (2011). Chinese characters elicit face-like N170 inversion effects. *Brain Cogn*.

Weise, H., Stahl, J., & Schweinberger, S. R. (2009). Configural processing of other-race faces is delayed but not decreased. *Biological Psychology, 81*(2), 103-109.

Xu, Y., Liu, J., & Kanwisher, N. (2005). The M170 is selective for faces, not for expertise. *Neuropsychologia, 43*(4), 588-597.

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

Zeki, S., Watson, J. D., Lueck, C. J., Friston, K. J., Kennard, C., & Frackowiak, R. S. (1991). A direct demonstration of functional specialization in human visual cortex. *J Neurosci, 11*(3), 641-649.

Zhao, L., & Bentin, S. (2008). Own- and other-race categorization of faces by race, gender, and age. *Psychon Bull Rev, 15*(6), 1093-1099.

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*(8), 1741-1749.

Acknowledgements

I would like to thank first and foremost, my supervisor Guillaume Thierry and his family. Thank you for taking the chance on me, and always being supportive. Everything I have worked for and have in my career has been influenced by you, and there is no possible way I could repay the kindness shown to me by you and your family. I am indebted to you, thank you so much.

To Jan-Rouke Kuipers, a superb colleague, office mate, and importantly, friend. I can only hope to have somebody like you as a colleague in the future. Thank you for keeping me sane when I needed it, and kicking me in the pants when I needed that too. I look forward to sailing with you again, and I hope to see you and your beautiful family soon!

To my lab colleagues past and present, Nicola Savill (one of the first people I met in Bangor, and a very special girl!), Bastien Bouttenet, Yanjing Wu, Noriko Hoshino, Eirini Sanoudaki, James Keidel, Dusuna Dorjee & Alice Foucart. Thank you all for making Bangor an excellent place to live and work!

I would also like to thank my other supervisors and collaborators, especially Stephan Boehm, Debra Mills, Clara Martin, Sancho Moro, & Marie Lallier, for their continued friendship and always helpful advice. Any other friends or colleagues that I have missed, don’t be upset, its just that this is my last day in Bangor and I do not have time to write this in full, I have to leave! I will miss all of you!

Finally, special thanks to my family for always being supportive.