The initial representation of individual faces in the right occipito-temporal cortex is holistic: Electrophysiological evidence from the composite face illusion

Corentin Jacques
Unité Cognition & Développement; Laboratoire de Neurophysiologie, Université Catholique de Louvain, Louvain-la-Neuve, Belgium

Bruno Rossion
Unité Cognition & Développement; Laboratoire de Neurophysiologie, Université Catholique de Louvain, Louvain-la-Neuve, Belgium

Identifying a facial feature (e.g. the eyes) is influenced by the position and identity of other features (e.g. the mouth) of the face, supporting the view that an individual face is represented as a whole in the human brain. To clarify how early in the time-course of face processing this holistic individual representation is accessed we recorded event-related potentials during an adaptation paradigm of the composite face illusion (CFI). Observers performed a matching task on top halves of two faces presented sequentially. For each face pair, top and bottom face halves could be both identical, both different, or only the bottom half differed. The signal was larger over the right occipito-temporal cortex at about 160 ms (N170) when the attended top half differed between the two faces than when identical top halves were repeated. Crucially, a larger N170 was also found when the top halves of the two faces were the same, yet the observers had the illusion that they differed (CFI). This effect was not found when the two face halves were spatially misaligned. These observations indicate that the earliest perceptual representation of an individual face in the human brain is holistic rather than based on independent face parts.

Keywords: face perception, holistic processing, event-related potential, N170, adaptation


Introduction

The speed at which complex visual stimuli (i.e. object, faces) are categorized in the human brain is a topic of high interest in Cognitive Neuroscience (Rossion & Jacques, 2008; Thorpe, Fize, & Marlot, 1996). Of particular interest is the question of when, and along which processing time-course, different features of a visual stimulus are integrated into a global representation of the stimulus. A human face is perhaps the most relevant visual stimulus to address this question given that it is composed of multiple cues such as external and internal features (eyes, ears, nose, mouth, etc.), the relative distances between these features (e.g. inter-ocular distance), reflectance properties of the skin surface (color and texture) and three-dimensional global shape of the face. It is generally acknowledged that these facial cues are not processed independently from each other but that the individual face is perceived as a gestalt, or a “holistic” representation (Farah, Wilson, Drain, & Tanaka, 1998; Galton, 1883; Rossion, 2008b; Sergent, 1984; Tanaka & Farah, 1993). Evidence that an individual face is processed holistically come from the demonstration that the perception of a facial feature is affected by alterations to the identity or the position of one or several other features of the face (e.g. Farah et al., 1998; Sergent, 1984; Tanaka & Farah, 1993; Young, Hellawell, & Hay, 1987). The most compelling illustration of this phenomenon comes from an adaptation of the so-called “composite face effect” (Young et al., 1987) to create a visual illusion (Rossion, 2008b) in which identical top halves of faces are perceived as being slightly different if they are aligned with different bottom halves (Figure 1). This composite face illusion (CFI) is a particularly clear demonstration that the features of a face (here the two halves of a single face) cannot be perceived independently from each other, but are rather integrated into an undecomposed whole representation.

However, it remains to be determined whether a face stimulus is first/initially represented as a collection of independent features which are later integrated into a holistic representation (i.e., a hierarchical mode of face representation, e.g., Haxby, Hoffman, & Gobbini, 2000; Jiang et al., 2006; Ullman, 2007), or if the initial representation of an individual face in the human brain is already global/holistic (e.g., Sergent, 1986).

Here we addressed this question by recording scalp event-related potentials (ERP) during an identity adaptation...
paradigm (Caharel, d'Arripe, Ramon, Jacques, & Rossion, 2009; Ewbank, Smith, Hancock, & Andrews, 2008; Harris & Nakayama, 2008; Jacques, d’Arripe, & Rossion, 2007; see also Kovacs et al., 2006 for a general adaptation paradigm to the face category) of the composite face illusion. Specifically, in each trial observers were presented with an adapting face composed of distinct top and bottom parts (i.e. a composite face) for 3 seconds, rapidly followed by a second (test) composite face. The adapting and test composite faces could either (1) have identical top and bottom face halves (same), (2) differ only in their bottom halves (bottom different), or (3) differ in both their top and bottom halves (top + bottom different) (Figure 1). In addition, adapting and test faces were either presented both in an aligned or both in a misaligned format, since it is known that spatial misalignment of the face halves disrupts holistic processing (Young et al., 1987; Figure 1). In line with previous adaptation studies, we expected that the presentation of 2 different composite faces (i.e. top + bottom different) would lead to a larger amplitude (release from adaptation) as early as the face-sensitive N170 component on occipito-temporal recording sites (Jacques et al., 2007). Most importantly, if the initial representation of individual faces is holistic (i.e. global), we expected a significant release from adaptation on the N170 when the fixated top parts of the adapting and test faces were identical but were perceived as different (i.e. bottom different condition in aligned format). If this effect takes place as early as individual face representations are accessed (i.e., N170, Jacques et al., 2007; Jacques & Rossion, 2006), this would support the view that the initial representation of an individual face in the human brain is inherently global/holistic. However, if the early sensitivity to facial identity at the level of the N170 is based on independent individual facial features, then we should not observe a larger release from identity adaptation in this condition (bottom different condition in aligned format) as early as the N170 as compared to misaligned faces. Such differential effects, if not observed at the N170, may take place rather at later time points during the course of face processing.

**Methods**

**Subjects**

Seventeen paid volunteers (5 men, 1 left-handed, mean age 21.6 ± 1.7 years) participated in the experiment. All participants had normal or corrected-to-normal vision.
Stimuli

Stimuli used in this experiment were constructed from 30 top parts of faces (17 males and 13 females) and 43 bottom parts of faces (23 males and 20 females). Face tops and bottoms were taken from whole faces cropped to remove background, hair and everything below the chin and horizontally cut in two halves above the nostrils. All resulting top and bottom parts were converted to grayscale and equated for mean pixel luminance. Composite face stimuli were then created by pairing each face top with 2 different bottom parts (some bottom parts being used for more than one top), resulting in 30 pairs of stimuli; each pair with an identical top part but different bottom parts (see Figure 1). The created pairs of composite faces were either formed by top and bottom face parts precisely aligned to each other (“aligned” format) or were formed by the very same top and bottom parts with the bottom slightly offset to the right (30 pixels = 0.6° of visual angle). These stimuli with offset bottom parts are commonly called “misaligned” composite faces. Aligned faces subtended 2.5° in width × 3.4 in height of visual angles, while misaligned faces subtended 3.1° in width × 3.4 in height (see Figure S1). A small gap of 3 pixels in height (<0.1°) between the top and the bottom parts was used so that participants could easily identify the border of the top and bottom parts in the aligned condition (e.g. Goffaux & Rossion, 2006).

Procedure

Stimuli were displayed against a light gray background using E-prime 1.1 (PST) at 110 cm viewing distance. In each trial, two face stimuli were presented sequentially, either both in the aligned or both in the misaligned format. Both stimuli were presented at the center of the screen and overlapped completely. A trial started with a fixation point displayed at the center of the screen for 200 ms (Figure 1). Two hundred milliseconds after the offset of the fixation point (randomized between 100 and 300 ms), the first face (adapting stimulus) appeared for about 2800 ms (2600 to 3000 ms) with the top part being presented at fixation. The offset of the adapting face was followed by an interval of random duration (150 to 350 ms), and then the second face (test face) for 200 ms. An inter-trial-interval of about 1600 ms (1500 to 1700 ms) separated the offset of the test face from the next trial. Identity of the top and the bottom parts were manipulated between the adapting and the test face in three conditions. In one third of the trials, both the top and bottom parts of the test face stimulus were different from those of the adapting face (Figure 1, bottom row—“top + bottom different” condition). Importantly, the same pairs of faces were presented in the aligned and misaligned orientations, so that any effect cannot be attributed to the particular pairing of adapting and test faces. Participants were instructed to attend only to the top part of the face (appearing at fixation) and to press one of two response keys (keys counterbalanced across subjects) corresponding to whether the adapting and test top parts were the same or different pictures. Previous studies have shown that participants maintain gaze fixation on the top part of the face according to the instructions, and nevertheless show a strong composite face effect in this condition (de Heering, Rossion, Turati, & Simion, 2008). Note that, according to these instructions, both the same and bottom different conditions required a “same” response and only the top + bottom different condition required a “different” response. Participants performed 90 trials per condition, resulting in 540 trials. Consequently, each top parts of faces was repeated 18 times as adapter and 18 times as test stimuli and each bottom face part was repeated either 12 or 18 times as adapter and either 12 or 18 times as test stimuli. The order of conditions was randomized within each block.

EEG recording

Scalp EEG was recorded from 58 tin electrodes mounted in an electrode cap (Quik-cap, Neuroscan Inc.), with a left earlobe reference. Two pairs of bipolar electrodes were used to record vertical and horizontal eye movements. Electrode impedances were kept below 10 kΩ. EEG analog signal was digitized at a 1000-Hz sampling rate and a digital anti-aliasing filter of 0.27 * sampling rate was applied at recording (at 1000 Hz sampling rate, the usable bandwidth is 0 to approximately 270 Hz). EEG data were analyzed using EEProbe 3.2 (ANT, Inc.) and Matlab. After filtering of the EEG with a digital 30-Hz low-pass filter, time windows in which the standard deviation of the EEG on any electrode within a sliding 200-ms time window exceeded 35 µV were marked as either EEG artifacts or blink artifacts. Blink artifacts were corrected by subtraction of a vertical electrooculogram (EOG) propagation factors based on EOG components derived from principal component analyses. Trials containing EEG artifacts were rejected. Participants’ averages were baseline corrected using the 100-ms prestimulus epoch and then re-referenced to a common average reference.

ERP analyses

Two separate sets of analyses were performed on the ERP signal recorded at the scalp in response to the second
face of the trial sequence (test face). First, we performed conventional analyses on specific visual potentials, as classically done in studies of face perception (e.g. Bentin, McCarthy, Perez, Puce, & Allison, 1996; Itier & Taylor, 2002; Rossion et al., 1999): the P1 (maximal at approximately 105 ms) and N170 (maximal at approximately 165 ms) components, recorded at posterior sites. Amplitude values of the P1 and the N170 were measured at five pairs of occipito-temporal electrodes in the left and right hemisphere where both components were the most prominent (P7/8, P5/6, PO7/8, PO5/6, O1/2). Amplitudes were quantified for each condition as the mean voltage measured within 30-ms windows centered on the grand average peak latencies of the components’ maximum. The amplitude values of each component were then submitted to separate repeated measures analyses of variance (ANOVA) with the factors alignment (aligned, misaligned), adaptation (same, bottom different, top + bottom different), hemisphere (left, right) and electrode (five levels). In ERP analyses for the bottom-different condition, we included data for both correct and incorrect responses and not only for incorrect responses for two reasons. First, incorrect trials were too few (mean = 16.5 trials; SD = 15.4 trials) to be able to get a reliable ERP waveform for each subject. Second, the composite illusion does not only result in incorrect responses, it also generates increased response times even when subjects provide a correct response (e.g., Rossion & Boremanse, 2008; Young et al., 1987; see Figure 2). It is therefore almost impossible to determine whether the illusion took place or not on any given trial based on behavioral data, and all trials may be relevant in the experiment.

A second set of analyses was carried out to characterize more precisely the time-course of face holistic processing in the human brain during this experiment. Specifically, we analyzed ERP data at each electrode as a function of time in a series of pair-wise comparisons between the different conditions. The first two comparisons were

![Figure 2](https://jov.arvojournals.org/pdfaccess.ashx?url=/data/journals/jov/933553/)  
**Figure 2.** Behavioral results for the matching task of composite faces. A. Error rates are shown on the left and response times on the right of the figure. Error bars are standard errors. B. Distribution of correct (full lines) and incorrect (dashed lines) response times in the composite task, separately for same and bottom-different conditions in the aligned and misaligned face format. The number of responses in successive 20 ms time-bins is plotted as a function of time from stimulus (test face) onset. Numbers in parenthesis represent respectively the mean and median response times over all correct trials. For aligned trials, the composite face illusion affected even the very early part of the distribution.
performed between the ERP elicited in the same condition and the ERP elicited in the top + bottom different condition, both for aligned and misaligned faces. Given that these comparisons were carried out between the condition where both the top and bottom of the face are repeated (same) and the condition where both the top and bottom of the face are not repeated (top + bottom different), we considered these comparisons to reflect an effect of adaptation to a whole individual face.

The next two comparisons were performed between the same condition and the bottom different condition, again both for aligned and misaligned faces. These comparisons were carried out between the condition where both the top and the bottom of the face are repeated (same) and the condition in which the identity of the top of the test face was perceived as different in the aligned condition only, as the result of the composite illusion (i.e. adaptation to identity in composite illusion). To statistically identify the onset latency of the differential ERP responses between the conditions compared, we performed a permutation test (see Blair & Karniski, 1993; Nichols & Holmes, 2002) on each scalp electrode and each time sample. This method was used previously to assess the time course of ERP face identity adaptation effects (Jacques et al., 2007). In a given permutation sample, the ERP data (consisting of the whole electrode * time-point matrix) for the two conditions compared are randomly permuted within each subject (i.e. paired comparisons) to obtain two new bins of size N. Because permutation shuffles the assignment of the conditions, the difference between the means of the two new bins reflects the difference between conditions under the null hypothesis. We performed 5000 of all $2^{17}$ possible permutations to generate a distribution of ERP differences under the null hypothesis. Comparing the observed ERP difference between the two conditions compared with the permutation distribution allowed estimating the probability that this observed ERP difference was due to chance (i.e. a p-value). The results of this analysis are displayed in time by electrode statistical plots in which significant differences between conditions compared are color-coded as a function of the amplitude of ERP difference (see Figure 6).

To minimize the probability of type I errors (false positives) due to the large number of comparisons performed, only significant differences at the $p < 0.01$ (two-tailed), that lasted for at least 20 consecutive time-samples and included a cluster of at least two neighbor electrodes were considered.

**Results**

**Behavioral results**

Error rates and response times at the face top part matching task (Figure 2) were submitted to repeated measures ANOVAs with the factors adaptation (3 levels) and alignment (2 levels).

The ANOVA performed on error rates revealed a significant effect of adaptation ($F(1.1, 18.2) = 9.1$, $p < 0.01$), alignment ($F(1, 16) = 12.6$, $p < 0.005$) and a significant interaction between the two factors ($F(1.1, 17.8) = 22.1$, $p < 0.0005$). As expected, participants made many more errors with aligned compared to misaligned faces in the bottom different condition ($p < 0.0001$). Moreover, participants made also more errors in misaligned compared to aligned faces in the top + bottom different condition ($p < 0.05$). Last, participants made more errors in the top + bottom different compared to same ($p < 0.05$), probably due to the oddity of the correct response in this condition (“different”).

The ANOVA performed on correct response times also revealed a significant effect of adaptation ($F(1.4, 22.7) = 14.6$, $p < 0.0005$), alignment ($F(1, 16) = 32.3$, $p < 0.0001$) and a significant interaction between the two factors ($F(1.2, 18.9) = 23.6$, $p < 0.0001$). The interaction revealed that participants were much slower for aligned compared to misaligned faces in the bottom different condition ($p < 0.0001$). They were also slightly slower in aligned compared to misaligned faces in the same condition ($p < 0.05$) and the opposite was observed in the top + bottom different ($p < 0.05$). In addition, they were generally slower in the top + bottom different condition ($p < 0.005$).

In summary, we observed a strong composite face effect (i.e. increased errors and response times for aligned compared to misaligned in the bottom different condition) in our matching task, replicating the results of numerous previous behavioral studies (e.g., Goffaux & Rossion, 2006; Le Grand, Mondloch, Maurer, & Brent, 2004; Rossion & Boremanse, 2008).

**Electrophysiological results**

**ERP components analyses**

Grand average ERP waveforms elicited by the test faces in the different conditions are depicted in Figure 3 for two occipito-temporal electrodes (PO7/PO8).

At the level of the P1 (see Figure 4), a significant effect of alignment ($F(1, 16) = 13.1$, $p < 0.005$), qualified by a significant interaction between alignment and hemisphere ($F(1, 16) = 16.6$, $p < 0.001$) revealed that the P1 was larger in the misaligned than in the aligned condition, but only in the left hemisphere ($p < 0.001$), not in the right hemisphere ($p = 0.43$). There was also a significant effect of adaptation ($F(1.8, 29.2) = 8.8$, $p < 0.005$), indicating a larger P1 in the top + bottom different condition relative to same ($p < 0.005$) and bottom different ($p < 0.005$) conditions. The P1 amplitude for these latter two conditions did not differ ($p = 0.57$). The effect of adaptation was slightly larger on the most posterior electrodes (PO7/8, PO5/6, O1/2) as revealed by a significant adaptation × electrode interaction ($F(4.5$, $p = 0.001$).
However, the adaptation effect was significant on all 5 pairs of electrodes analyzed. Importantly, the adaptation × alignment interaction was not significant ($F(1.6, 25) = 0.8, p = 0.43$), indicating that the adaptation effect did not differ between aligned and misaligned faces at the P1 level.

Misaligned test faces elicited a larger N170 than aligned faces ($F(1, 16) = 24.5, p < 0.005$) (see Letourneau & Mitchell, 2008). The main effect of adaptation was not significant ($F(1.6, 26.2) = 2.5, p = 0.1$) but, importantly, a significant interaction between alignment and adaptation ($F(1.5, 24.1) = 4.1, p < 0.05$) revealed that the effect of
The time course of individual face holistic encoding

Point-by-point permutation tests were performed to compare same to top + bottom different conditions and same to bottom-different conditions in aligned and misaligned format. The differences between conditions compared are plotted as subtraction waveforms in Figure 5 (left part) for 58 electrodes, with 3 occipito-temporal electrodes in each hemisphere being highlighted.

Time by electrodes plots of the significant identity adaptation effects were observed at around 160 ms in the aligned format and slightly later (185 ms) and smaller in the misaligned format. Interestingly, these differences were both restricted to occipito-temporal electrodes in the right hemisphere, with concomitant difference of opposite polarity at anterior centro-temporal sites in the left hemisphere (Figures 5 and 6). This early adaptation effect lasted until about 280 ms both for the aligned and misaligned conditions.

Time by electrodes plots of the significant identity adaptation effect in the composite illusion (i.e. same vs. top + bottom different) are shown in the right column of Figure 6. Consistent identity adaptation effects were observed at around 160 ms in the aligned format and slightly later (185 ms) and smaller in the misaligned format. Interestingly, these differences were both restricted to occipito-temporal electrodes in the right hemisphere, with concomitant difference of opposite polarity at anterior centro-temporal sites in the left hemisphere (Figures 5 and 6). This early adaptation effect lasted until about 280 ms both for the aligned and misaligned conditions.

To summarize, the N170 was larger for bottom-different adaptation was significant in the aligned format ($p < 0.01$) and not in the misaligned format ($p = 0.52$) (Figure 3). An additional ANOVA performed only on the N170 recorded in the aligned format revealed a significant interaction between adaptation and hemisphere ($F(1.7, 26.9) = 4.2$, $p < 0.05$). In the right hemisphere, the N170 was larger for bottom different compared to same ($p < 0.01$), and for top + bottom different compared to same ($p < 0.005$). The N170 elicited by top + bottom different was only marginally larger than for bottom different ($p = 0.07$). In the left hemisphere, the only marginally significant effect was a larger N170 for top + bottom different compared to same ($p = 0.09$), the other two comparisons being non-significant ($p$'s $> 0.11$). In addition to the analyses performed on the N170 elicited by correct and incorrect response trials (see Methods), we performed an ANOVA on the N170 elicited by correct trials only, comparing same to bottom-different conditions in the aligned format. Similar to the when the analyses was performed on correct and incorrect trials, this ANOVA revealed a significantly larger N170 for the bottom-different compared to the same condition ($F(1,16) = 10.3$, $p < 0.01$). The adaptation × hemisphere was no longer significant ($F(1,16) = 2.2$, $p = 0.16$) even though the adaptation effect was still much more reliable in the right ($p = 0.003$) than in the left ($p = 0.04$) hemisphere.

To summarize, the N170 was larger for bottom-different and top + bottom different compared to the same condition. This effect was only observed in the aligned and not in the misaligned condition, in line with the composite effect observed on behavioral data, and was found only over right hemisphere electrodes. In other words, the differential N170 amplitude between bottom-different and same conditions is consistent with the perceptual illusion (i.e. top parts of adapting and test faces are perceived as different even though only the bottom parts are different) that occurs only in the aligned format.
These early differences were due to a slight latency difference between same and bottom different conditions at the level of the P1 peak, creating artificial amplitude differences in the N170 slope (slightly larger amplitude for same condition). An additional point-by-point analysis indicated that there was no longer significant difference between these conditions in the 120–140 ms time-window after correcting for the P1 latency difference (see Appendix A). In contrast, the difference during the N170 time-window remained conspicuous.

When considering later ERP responses (>300 ms post-stimulus), large differences were found between 350 ms and 450 ms, distributed over bilateral occipito-temporal regions with a polarity reversal over fronto-central electrodes. These differences were observed on all comparisons except in the misaligned format when comparing same to bottom different conditions (Figure 6). In this last comparison, the only significant difference was observed from 440 to 520 ms over centro-parietal electrodes. A last strong difference started slightly before 500 ms only when comparing same to top + bottom different conditions. This effect is due to the presence of a large ERP component, maximal over parietal regions, in the ERP response to the top + bottom different condition (Figure S2). This component is most probably a P300 component (see Polich, 2007 for a review) generated due to the fact that these conditions (aligned and misaligned) required a “different” response which had a smaller probability than “same” response (1/3 vs. 2/3 of the trials respectively) in the present study.

**Discussion**

Replicating recent observations (Caharel et al., 2009; Ewbank et al., 2008; Jacques et al., 2007), we found that
when an individual face is different from a previously presented adapter face, there is a release from adaptation starting at around 160 ms post-stimulus onset, in the right occipito-temporal cortex. That is, the N170 was larger in amplitude when pairs of different compared to repeated faces were presented, an effect that is also observed in some face identity repetition studies (e.g., Campanella et al., 2000; Heisz, Watter, & Shedden, 2006; Itier & Taylor, 2002; Jemel, Pisani, Calabria, Crommelinck, & Bruyer, 2003; but see e.g. Bentin & McCarthy, 1994; Schweinberger, Huddy, & Burton, 2004), but appears to be most consistent and large in face identity adaptation paradigms with a long presentation duration of the adapter face and a short inter-stimulus-interval, as used here (Caharel et al., 2009; Jacques et al., 2007). This observation indicates that different individual faces can be discriminated reliably in the human visual cortex as early as during the time-window of the so-called N170 ERP component (Bentin et al., 1996; Jeffreys, 1996; Rossion & Jacques, 2008 for a review). Unlike earlier and spurious differences that may arise from the repetition of physically identical images (i.e. at the level of the visual P1 component), this sensitivity to individual faces at 160 ms cannot be attributed to mere image-based differences, since it is abolished by flipping
the exact same stimuli upside-down (Jacques et al., 2007), and is also found even when there is a substantial (30°) change of viewpoint between the adapter and the test face (Caharel et al., 2009). Here, interestingly, we found that the release from adaptation started slightly later (at about 185 ms) when the two halves of the faces were spatially misaligned, suggesting that disrupting the whole facial stimulus delays the coding of individual faces (as upside-down inversion does, see Jacques et al., 2007). Together with the finding that inverting faces abolishes the release from adaptation effect on the N170 (Jacques et al., 2007), this absence of a release from adaptation on the N170 for the top + bottom different misaligned condition suggests that the effects observed on the N170 in the aligned condition (i.e. same vs. bottom-different and same vs. top + bottom different) are not simply due to the magnitude of the physical change between adapting and test faces.

The novel and most important observation of the present study is that when observers erroneously perceived two physically identical top parts of faces as being slightly different, increasing their mistakes and response times (the hallmark of the behavioral composite face illusion), there was also a release from adaptation in the electrophysiological response starting at about 165 ms, in the right hemisphere. This observation provides compelling evidence that as soon as the percept of the test face is sufficiently detailed to allow discriminating it from another individual face (the adapter), the parts that make-up the individual face stimulus are integrated into a global (i.e., holistic) representation. That is, the initial representation of an individual face in the human brain appears to be inherently holistic. Importantly, when the two halves of the face were spatially misaligned and thus the top part of the faces were correctly perceived as being identical, there was no release from adaptation, even though the bottom parts differed between the adapting and the test stimuli. In short, here we clarified the time-course of the composite face illusion in the human brain, a well-known phenomenon in the behavioral face processing literature which is taken as the strongest evidence that the facial features of an individual face are integrated into a whole face representation (Maurer, Le Grand, & Mondloch, 2002; Rossion, 2008b; Young et al., 1987).

It is important to make a distinction between our findings and earlier reports of a sensitivity to the whole face stimulus at the level of the N170 (e.g., George, Jemel, Fiori, Chaby, & Renault, 2005; Latinus & Taylor, 2005; Letourneau & Mitchell, 2008). The behavioral literature indicates that faces are processed holistically both at the basic category level (i.e. categorizing the visual stimulus as “a face”, or detecting a face in a visual scene) and at the individual level (i.e. categorizing the stimulus as an individual face). Evidence that basic-level categorization of a visual stimulus as a face relies on holistic processing comes from the finding that subjects are slower to detect a face in a visual scene when the first-order organization of face is disrupted by upside-down inversion (Rousselet, Mace, & Fabre-Thorpe, 2003). Further evidence comes from the observation that, in certain circumstances basic-level categorization of the face depends heavily on a global representation, for instance when the observer perceives a face in two-tones “Mooney” face stimuli (Mooney, 1957; Moore & Cavanagh, 1998), or in a painting of the artist Arcimboldo (Hulten, 1987) which are composed on non-facial features. ERP studies have provided evidence that holistic processing of faces at the basic level takes place during the N170 time-window, since the amplitude of the N170 is larger when observers correctly perceive Mooney face stimuli (George et al., 2005; Latinus & Taylor, 2005) or faces in Arcimboldo paintings (Caharel et al., unpublished data; see Rossion & Jacques, 2008). Also, as observed here and in a recent study (Letourneau & Mitchell, 2008), breaking a single face stimulus in two separate parts affects the N170 amplitude, suggesting that the global first-order organization of the stimulus is coded at the level of the N170. However, this effect leads to a somewhat paradoxical increase of the N170 amplitude (Letourneau & Mitchell, 2008; the present study), similarly to the effect of face inversion (e.g. Itier & Taylor, 2002; Rossion et al., 1999), or to the presentation of isolated eyes on the N170 (Bentin et al., 1996; Itier, Alain, Sedore, & McIntosh, 2007). The meaning of this higher ERP amplitude to simple face stimulus transformations (usually associated with a delayed latency) is currently debated (see e.g. Bentin et al., 1996; Itier et al., 2007; Jacques et al., 2007; Rossion et al., 1999) and remains unclear.

Here, in contrast to these latter studies investigating basic-level face categorization, we used the composite face illusion in conjunction with adaptation to clarify when an individual holistic face representation is activated in the visual cortex. At the behavioral level, holistic processing of individual faces is revealed by the finding that the perception of the identity of a feature is strongly influenced by the identity or the position of the other features of a face, such as in the “whole-part advantage” paradigm (Tanaka & Farah, 1993) or in the CFI (Goffaux & Rossion, 2006; Young et al., 1987; the present experiment). The present results thus indicate that populations of neurons in the right occipito-temporal cortex that participate in generating the N170 component code the two face halves as an undissociated representation of the whole individual face.

The observation that the CFI takes place at around 160 ms in the N170 time-window poses serious constraints on the time-course and unfolding of face processes in the human brain. To understand their significance, these observations should be put in the general context of the time-course of face processing in the human brain. The earliest visual face-sensitive responses can be observed as early as 100 ms on the human scalp (P1 component, or M1 in MEG studies, e.g. Halgren, Raji, Marinkovic, Jousmäki, & Hari, 2000; Herrmann, Ehlis, Ellgring, &
However, these early P1/M1 effects appear to be related to low-level visual differences between faces and objects (e.g., differences in spatial frequency spectra, color distribution, global or local contrast) rather than to the activation of face representations in the human brain (Rossion & Jacques, 2008; Rousselet, Husk, Bennett, & Sekuler, 2008; Tanskanen, Nasanen, Montez, Paallysaho, & Hari, 2005). Solid evidence for activation of high-level face representations rather point to a time onset of about 120–130 ms (N170 time-window; Rossion & Jacques, 2008; Rousselet et al., 2008), a timing that is compatible with the average onset latency of face-selective cells in the monkey inferior-temporal cortex (at around 80–100 ms; Kiani, Esteky, & Tanaka, 2005), taking into account that these neurons should fire slightly later in the bigger human brain (see Foxe & Schroeder, 2005). This activation of facial representations during the N170 is based on the presence of clearly identifiable facial features and/or a first-order facial organization (Rossion & Jacques, 2008). As indicated above, the face representation starts to be sufficiently refined at about 160 ms to allow the discrimination between different individuals (Caharel et al., 2009; Ebanks et al., 2008; Jacques et al., 2007; Jacques & Rossion, 2006).

Thus, given that the N170 represents an early stage of face perceptual encoding (Bentin et al., 1996; Eimer, 2000; Rossion & Jacques, 2008; Rousselet et al., 2008), it confirms that holistic processing of individual faces occurs at a perceptual level, in contrast to the recent suggestions that behavioral effects in the composite face paradigm have a decisional rather than a perceptual locus (Richler, Gauthier, Wenger, & Palmeri, 2008). While we do not question the fact that the composite visual illusion leads to decisional response biases (as found also here, see Figure 2), our observations with an ERP method allowing measuring directly the time-course of face processing nevertheless point to the perceptual origin of the effect. Most importantly, these observations suggest that as soon as the perceptual system is sensitive to the individuality of the face (~160 ms), the individual features are processed in a holistic representation. That is, in the bottom-different aligned condition, the top part of the face is physically identical between the adapter and the test face, but it is perceived as different as early as the system is sensitive to individual faces. To put it differently, the holistic representation of an individual face is not accessed after an analysis of the local features of the face that would then be integrated into such a global representation. Hence, these observations are difficult to reconcile with a hierarchical process in which facial parts or fragments would be extracted first and then combined into a global representation of the individual face (e.g., Harris & Nakayama, 2008; Haxby et al., 2000; Jiang et al., 2006; Ullman, 2007). Rather, our observations suggest that, at the time window of the N170, the individual test face is processed as a whole and compared to an internal representation built previously from the adapter face, also stored as a holistic representation (a global template). Such a holistic representation contains information about the local shape of the features as well as their positions with respect to each other and to the global contour over the entire face (e.g., Barton, Zhao, & Keenan, 2003; Sergent, 1984; Young et al., 1987). Thus, changing the bottom half of the face between the adapter and the test modifies the perception of the whole face (Young et al., 1987; see Figure 1). Consequently, it is the whole test face, including the top half, which does not match the encoded template, leading to a release from adaptation in this condition. We also note that this effect, which emerges significantly at the level of the N170 and is prolonged during a longer time-window, is not as large as when the two halves of the face differ physically, since the whole faces differ even more largely in that condition.

Importantly, our results and this interpretation do not mean that local internal facial features such as eyes, eyebrows or nose, or larger fragments (Harel, Ullman, Epshtein, & Bentin, 2007), would not be part of the face representation activated at the time of the N170 visual evoked potential. There is clear evidence that facial features modulate the electrophysiological signal within the time window of the N170 (e.g., Bentin et al., 1996; Eimer, 2000; Harris & Nakayama, 2008; Itier et al., 2007; Schyns, Jentzsch, Johnson, Schweinberger, & Gosselin, 2003; Zion-Golumbic & Bentin, 2007). Moreover, local facial features, and not only their global organization, are an important part of what defines an individual face, even in a holistic representation (e.g., Wallis, Siebeck, Swann, Blanz, & Bulthoff, 2008). However, what our results suggest is that the perception of a feature (here in the top half) in an individual face depends, from the outset, on the perception of the other feature(s) (here in the bottom half) that make the whole individual face. In other words, while individual local features are undoubtedly highly diagnostic of face identity, they do not appear to be represented initially as separate entities from the whole face. This proposal is compatible with a strong view of holistic face representation as advocated by Tanaka, Farah and colleagues (Farah et al., 1998; Tanaka & Farah, 1993) according to which faces are represented only holistically, or at least that the holistic representation precedes and overrules the representation of facial parts.

This proposal of an initial holistic representation of the individual face is also in tune with the view that face processing evolves from a global coarse representation to a representation being progressively refined with finer resolution information, i.e., a coarse-to-fine mode of face perception (Sergent, 1986; Sugase, Yamane, Ueno, & Kawano, 1999). This coarse-to-fine mode of processing face stimuli is compatible with more general models of visual processing in which the perception or the awareness of the global shape/organization of an object or a visual scene precedes the analysis of local visual information (e.g., Flavell & Draguns, 1957; Hochstein & Ahissar, 2002;
Navon, 1977). One of the common ground between these models is that the analysis of the global information may provide a first approximation of what the object or the scene is, and guide a more detailed visual scrutiny. For instance, in the “reversed hierarchy theory” (Hochstein & Ahissar, 2002), conscious perception of the global shape first occurs in high-level visual areas containing neurons with larger receptive fields. Finer-grained representations could then be obtained through reentrant connections with lower-level visual areas containing neurons with smaller receptive field and capable or representing more local information (see also Lamme & Roelfsema, 2000; Mumford, 1992). Such a coarse-to-fine mode of processing, with an initial activation in high-level visual areas sensitive to faces and reentrant connections with lower visual areas to refine the representation, would also be suitable with the processing of an individual face in the human brain (Rossion, 2008a). More precisely, there is evidence from fMRI adaptation studies that among the face-sensitive visual areas of the occipito-temporal cortex (see Haxby et al., 2000; Sergent & Signoret, 1992), the higher level “fusiform face area (“FFA”, Kanwisher, McDermott, & Chun, 1997) in the right hemisphere encode faces holistically (Harris & Aguirre, 2008; Rossion et al., 2000; Schiltz & Rossion, 2006). Most interestingly, holistic representation of individual faces in the right ‘FFA’ was recently found using a composite face illusion in fMRI, both in a block design (Schiltz & Rossion, 2006) and an event-related paradigm similar to the present study (Schiltz et al., submitted). Along these lines, it should be noted also that in the present study the early release from adaptation due to the integration of features into a holistic individual face representation was found only over the right occipito-temporal scalp region. These observations are compatible with the well-known dominance of the right hemisphere in processing faces in general (e.g. Sergent & Signoret, 1992), and holistic face representations in particular (e.g., Hillger & Koenig, 1991). Besides the FFA, the face-sensitive area in the inferior occipital gyrus (the ‘occipital face area’; Gauthier et al., 2000) also encodes face holistically as revealed by the CFI (Schiltz & Rossion, 2006), although less consistently than in the ‘FFA’, and might therefore also contribute to the effect observed in the present study.

In conclusion, the present study provides evidence that as early as an individual face is encoded in the right hemisphere at about 160 ms following stimulus onset, the representation of that face is holistic rather than based on independent features.

Appendix A

To ensure that the adaptation effect observed in the composite illusion (same vs. bottom different in aligned format) during the N170 time-window was not due to the latency difference at the level of the P1 peak and creating artificial differences in the 120–140 ms window, we performed the point-by-point analysis on the ERP data corrected for this latency difference. Specifically, we first measured the grand-average P1 latency in each condition based on the global field power (i.e. the standard deviation across all electrodes at each time-point; Lehmann &
Skrandies, 1980) and then re-aligned the ERP waveforms of each subject based on the difference between the P1 latency averaged across conditions and the P1 latency for a given condition. In the four comparisons described above, only the comparison between same and bottom different conditions in the align format was affected by this latency correction. That is, the P1 measured on the global field power was 1 ms earlier in the bottom different than in the same condition. We then performed the point-by-point analysis on the latency corrected ERP waveforms for the same vs. bottom different comparison only. In this comparison, after correcting for the P1 latency there was no longer significant difference between conditions in the 120–140 ms time-window (Figure A1). In contrast, the difference during the N170 time-window remained conspicuous.

**Acknowledgments**

Corentin Jacques and Bruno Rossion are supported by the Belgian National Fund for Scientific Research (Fonds de la Recherche Scientifique—FNRS). This work was supported by an ARC (Actions de Recherche Concertées) grant 07/12-007, Communauté Française de Belgique.

Commercial relationships: none.

Corresponding author: Corentin Jacques.

Email: corentin.g.jacques@uclouvain.be.

Address: Université Catholique de Louvain (UCL), Faculté de Psychologie et des Sciences de l’Education (PSP), 10, Place du Cardinal Mercier, 1348 Louvain-la-Neuve, Belgium.

**References**


human fusiform face area as determined by MEG. Cerebral Cortex, 10, 69–81. [PubMed] [Article]


