

# Time course and robustness of ERP object and face differences

**Guillaume A. Rousselet**

Centre for Cognitive Neuroimaging (CCNi) and  
Department of Psychology, University of Glasgow,  
Glasgow, UK



**Jesse S. Husk**

Department of Psychology, Neuroscience, & Behaviour,  
McMaster University, Hamilton, ON, Canada



**Patrick J. Bennett**

Department of Psychology, Neuroscience, & Behaviour,  
McMaster University, Hamilton, ON, Canada



**Allison B. Sekuler**

Department of Psychology, Neuroscience, & Behaviour,  
McMaster University, Hamilton, ON, Canada



Conflicting results have been reported about the earliest “true” ERP differences related to face processing, with the bulk of the literature focusing on the signal in the first 200 ms after stimulus onset. Part of the discrepancy might be explained by uncontrolled low-level differences between images used to assess the timing of face processing. In the present experiment, we used a set of faces, houses, and noise textures with identical amplitude spectra to equate energy in each spatial frequency band. The timing of face processing was evaluated using face–house and face–noise contrasts, as well as upright–inverted stimulus contrasts. ERP differences were evaluated systematically at all electrodes, across subjects, and in each subject individually, using trimmed means and bootstrap tests. Different strategies were employed to assess the robustness of ERP differential activities in individual subjects and group comparisons. We report results showing that the most conspicuous and reliable effects were systematically observed in the N170 latency range, starting at about 130–150 ms after stimulus onset.

Keywords: ERP, EEG, object processing, face processing, noise, textures, robust statistics, bootstrap, trimmed mean

Citation: Rousselet, G. A., Husk, J. S., Bennett, P. J., & Sekuler, A. B. (2008). Time course and robustness of ERP object and face differences. *Journal of Vision*, 8(12):3, 1–18, <http://journalofvision.org/8/12/3/>, doi:10.1167/8.12.3.

## Introduction

In the study of face processing, event-related potentials (ERPs) constitute a key technique, whereby single-trial EEG segments are averaged time-locked to stimulus onset. The literature on face ERPs has focused primarily on the N170, an occipital-temporal component that is systematically larger to faces compared to a broad variety of objects (Bötzel, Schulze, & Stodieck, 1995; Carmel & Bentin, 2002; Eimer, 2000a; Itier & Taylor, 2004; Rossion et al., 2000; Rousselet, Macé, & Fabre-Thorpe, 2004), as well as its positive counterpart, the VPP or the P150 (Jeffreys, 1996; Joyce & Rossion, 2005; Schendan, Ganis, & Kutas, 1998). The N170 spans a time window of about 130–200 ms, providing a coarse estimate of the time necessary to extract visual information about faces and other objects, at least at the categorical level. However, the focus on the N170 as the first marker of face processing has been challenged by several reports suggesting that there are earlier differences between face and object ERPs, circa 80–130 ms post-stimulus (Debruille,

Guillem, & Renault, 1998; Halit, de Haan, & Johnson, 2000; Herrmann, Ehlis, Ellgring, & Fallgatter, 2005; Itier & Taylor, 2002, 2004; Linkenkaer-Hansen et al., 1998; Rousselet, Macé, Thorpe, & Fabre-Thorpe, 2007). Some authors have even claimed that faces might be processed in as little as 50–80 ms after stimulus onset (George, Jemel, Fiori, & Renault, 1997; Mouchetant-Rostaing, Giard, Bentin, Aguera, & Pernier, 2000; Mouchetant-Rostaing, Giard, Delpuech, Echallier, & Pernier, 2000; Seeck et al., 1997). In monkey studies, differences between faces and objects have been reported typically in the range 80–120 ms (Logothetis & Sheinberg, 1996; Kiani, Esteky, & Tanaka, 2005).

The discrepancies in the reported onset of neural correlates of face processing might be due to differences in task demands, the presence of uncontrolled physical differences across stimuli, or the robustness of the statistical design (for example, individual differences may be obscured by averaging to differing degrees across studies).

Early ERP differences for faces could be the result of uncontrolled low-level differences across object categories

(Johnson & Olshausen, 2003; Rousselet, Macé, Thorpe, & Fabre-Thorpe, 2007; VanRullen & Thorpe, 2001). For instance, stimulus characteristics such as overall luminance, contrast, or spatial frequency and orientation components have not been equated across object categories. It has been suggested that the N170 varies with spatial frequency content of the stimulus (Goffaux, Gauthier, & Rossion, 2003; see also MEG results from Tanskanen, Näsänen, Montez, Päälyysaho, & Hari, 2005), and it is equally likely that differences in early ERP components could be due to low-level differences in stimuli rather than indicating differences in category-related processing *per se*. This is not to say that higher level and task factors do not influence ERP components. Indeed, top-down factors do influence early visual evoked responses (Bentin & Golland, 2002; Luck, Woodman, Vogel, 2000). However, the brain responses to various foveated stimuli like objects, words, animal, and human faces are hardly modulated by task factors before 200 ms after stimulus onset in a large majority of studies, as demonstrated by surface EEG (Carmel & Bentin, 2002; Cauquil, Edmonds, & Taylor, 2000; Lueschow et al., 2004; Schendan et al., 1998; Rousselet et al., 2004; Rousselet, Macé, et al., 2007; but see two exceptions in Eimer, 2000a, 2000b), intracranial recordings (Nobre, Allison, & McCarthy, 1998; Puce, Allison, & McCarthy,

1999), and MEG (Furey et al., 2006, Lueschow et al., 2004). In the present study, we used well-controlled stimuli (faces, houses, and noise textures, presented upright and inverted, Figure 1), with identical amplitude spectra, to evaluate the timing of face and object processing. Because form information is largely carried by phase rather than amplitude (Oppenheim & Lim, 1981; Sekuler & Bennett, 1996), individual houses and faces remained easily discriminable after this manipulation. However, this manipulation ensured that any difference in the EEG was not simply a function of differences in luminance or contrast or in the relative strength of specific frequency or orientation components across categories. Despite the important ERP literature on object and face processing, very few studies have used stimuli with equated amplitude spectra: (1) some of these studies have used noise textures (e.g., Allison, Puce, Spencer, & McCarthy, 1999; Jacques & Rossion, 2004; see also MEG data in Tanskanen, Näsänen, Ojanpää, & Hari, 2007); (2) and to the best of our knowledge only our previous study used objects like houses (Rousselet, Husk, Bennett, & Sekuler, 2005; Rousselet, Husk, Bennett, & Sekuler, 2007).

In addition to problems arising from uncontrolled image properties described above, the evaluation of face processing speed might be potentially affected by data analysis procedures. Most ERP experiments have relied



Figure 1. Examples of stimuli used in the experiment.

on the mean as a measure of central tendency (also called measure of location) and on group statistics to compare those means. There are many problems associated with both using the mean and classic parametric group statistics to analyze ERPs. To summarize, the mean is a good representation of data only if the data are normally distributed and is very sensitive to even small departure from normality (Wilcox, 2005). In other words, if the data are not normally distributed, and there is no a priori reason to believe that EEG data are, then the mean fails to portray the behavior of most of the individual trials. In the presence of outliers, the trimmed mean often provides a better estimate of central tendency. Following the work of Wilcox, we decided to use 20% trimmed means in the present paper (Wilcox, 2005; Wilcox & Keselman, 2003). Trimmed means were first compared at the group level, as is routinely performed in this field of research. However, drawing conclusions from group statistics can be problematic because effects can be driven by a few subjects, and alternatively, interesting effects may be present in single subjects but masked by group statistics. These limitations can be overcome by conducting single-trial analyses for each subject to estimate the number of subjects showing a particular effect. That way, the importance of early categorical differences can be weighted by the number of subjects actually showing them. In this paper we used both group and individual subject statistics to estimate the time course of face and object processing. These analyses were based on bootstrap statistics, which do not make strong assumptions about the underlying distribution of the data (Wilcox, 2005). We also propose simple strategies to apply this approach at the single-trial level. Specifically, instead of providing a binary output as to whether an effect is significant or not, we estimate the long run probability of reproducing that effect. Altogether, this hierarchy of data analyses provides a very useful tool to evaluate the robustness of an effect.

In the ERP literature, the speed of face processing has been estimated using two main types of comparisons, one relying on the time at which ERP to faces and other objects start to diverge significantly and one on the latency of the inversion effect. When faces are presented upside-down, the N170 is often delayed and larger compared to upright faces (Itier, Latinus, & Taylor, 2006; Jacques & Rossion, 2007; Rossion & Gauthier, 2002; Rousselet et al., 2004). A weaker and less consistent inversion effect has been reported for stimuli other than human faces in some studies (Eimer, 2000a; Itier et al., 2006; Rossion, Joyce, Cottrell, & Tarr, 2003; Rousselet et al., 2004). Inversion effects on the P1 have been reported for faces only, providing tentative evidence for early face specific mechanisms circa 100–130 ms (Itier & Taylor, 2002, 2004; Linkenkaer-Hansen et al., 1998). In the present study, category differences and the inversion effect were evaluated systematically at each time point with our new set of stimuli. We conclude that the first reliable category and inversion effects occur in the time

window of the N170, starting at about 130–150 ms after stimulus onset. Different single-trial analyses of the same data have been published previously (Rousselet et al., 2007).

## Methods

### Participants

Sixteen subjects participated in this experiment (age range 19–39, mean age 25). All subjects gave written informed consent and had normal or corrected-to-normal vision. Except for 4 members of the laboratory, subjects received \$10/hour for their participation. Twelve subjects were right handed, and 9 were female.

### Stimuli

We used front-view grayscale photographs of 10 faces and 10 houses centered in a  $5.2 \times 5.2$  deg background of average luminance (Figure 1; for more stimulus details, see Gold, Bennett, & Sekuler, 1999a; Rousselet et al., 2005; Husk, Bennett, & Sekuler, 2007). The house set was designed to have within-set homogeneity that is typical of faces, with key features found in the same relative positions across exemplars. A consistent configuration was used with houses because typical stimulus sets contain much more variability within houses (different number of windows, completely different layouts, etc.) than within faces (hair vs. no hair, different views and sizes, etc.), which can provide cues for discriminating between categories. All the stimuli were equated in terms of spatial frequency content by taking the average of the amplitude spectra of all 20 stimuli and then combining that average spectrum and the original phase spectra to reconstruct each individual stimulus. Ten textures were created by combining the mean amplitude spectrum across faces and houses with random phase spectra sampled from a Gaussian distribution. Thus, the noise textures had the same amplitude spectrum as the faces and houses (approximately pink noise). Finally, all stimuli were windowed by a 2D Gaussian function (with a width at half-height of the Gaussian window of 150 pixels for the original  $256 \times 256$  pixel images) to minimize edge effects when the stimuli were inserted in the full screen gray background (luminance =  $25.4 \text{ cd/m}^2$ ,  $23.5 \text{ deg} \times 30.1 \text{ deg}$ ).

### Experimental design

Subjects sat in a dimly lit sound-attenuated booth. Viewing distance was maintained at 70 cm by the use of a chinrest. Stimuli were presented for 80 ms (6 frames at

75 Hz) on a Sony Trinitron GDM-F520 monitor (resolution  $1024 \times 768$  pixels). Subjects were given 1080 ms to respond by pressing one of three keys, using three fingers from their dominant hand, to indicate whether a face, a house or a texture appeared on the screen. The button/category association was randomized for each subject. An experiment was composed of 12 blocks of 120 trials (1440 trials in total with 240 trials per condition). Within each block, there were two presentations of each item (face, house, texture) in two orientations (upright, inverted). On each trial, a blank screen was presented for about 200 ms, followed by a small black fixation cross (a 0.3 deg “+” in the middle of the screen) for a random duration ranging from 500 to 900 ms. Then a stimulus was presented for 80 ms, followed by a blank screen for 1000 ms during which time subjects were allowed to make a response to the categorization task (face, house or texture). After that delay, responses were considered incorrect. Trial durations thus ranged from 1780 to 2180 ms.

## EEG recording and analysis

### EEG recording

EEG data were acquired with a 256-channel Geodesic Sensor Net (Electrical Geodesics Inc., Eugene, Oregon; Tucker, 1993). Analog signal was digitized at 500 Hz and band-pass filtered between 0.1 and 200 Hz. The ground electrode was along the midline, anterior to Fz, and impedances were kept below 50 k $\Omega$ . Subjects were asked to minimize blinking, head movement, and swallowing.

### EEG preprocessing

EEG analysis was performed using EEGLAB (Delorme & Makeig, 2004) and Matlab. EEG data were referenced on-line to electrode Cz and re-referenced off-line to an average reference. Out of the original 256 electrodes, 60 electrodes were rejected by default in all subjects because they tended to be particularly noisy: the last row of posterior electrodes, electrodes around the ears and on the cheeks. Further electrodes were rejected on a subject-by-subject basis, leading to a minimum number of 179 electrodes, with a mean of 191 electrodes, still providing a very good coverage of the scalp. The signal was low-pass filtered at 30 Hz. Baseline correction was performed using the 300 ms of pre-stimulus activity. Artifacts were rejected based on absolute abnormal values larger than 120  $\mu$ V, and the presence of a trend, with an absolute slope of the linear trend larger than 75  $\mu$ V per epoch and a regression R-square value larger than 0.3. Only correct trials were averaged, using an interval from  $-300$  ms to  $+400$  ms. Across subjects and conditions, the minimum number of trials was 62, the maximum 231, and the mean 168.

### Measure of central tendency

Our measure of central tendency (or location) was the 20% trimmed mean. The trimmed mean was calculated at each electrode and each time point independently, by sorting the single-trials, trimming the lowest 20% of the distribution and the highest 20% of the distribution and then averaging the remaining trials (Wilcox, 2005, p. 56). The trimmed mean, by focusing on the bulk of the data in the central part of the single-trial distribution, provides a robust estimate of the location of a distribution. The 20% cutoff has proven a very good default in many situations (Wilcox, 2005; Wilcox & Keselman, 2003), especially when the comparison between measures of location relies on a percentile bootstrap (described below). Because the trimmed mean is not classically used in the ERP literature, two points should be stressed here. First, the median, often used to describe RT distributions, is nothing but the most extreme type of trimmed mean, where all data points are trimmed except one. Second, the mean, which is typically used as a measure of location in the ERP literature, is only a special case of a large class of location measures (Wilcox, 2005, p. 30). The mean is very sensitive to the tails of a distribution and thus can fail to represent the behavior of most of the elements of a distribution. In a nutshell, the mean should only be used under strict normality, a situation not often met with EEG data.

In addition to estimating the central tendency of the distribution at each time point and electrode, we used a more global description of the data, namely the standard deviation across electrodes computed at each time point (Lehmann & Skrandies, 1980). We call this measure global field amplitude (GFA). It corresponds exactly to what has been called global field power (GFP) in other publications (e.g., Rousselet et al., 2005). However, we think that GFP is a misleading term because power traditionally refers to the variance of the signal. GFA provides a compact description of the signal for each category, summarizing all electrodes in one vector. Because of the dipolar nature of early visual evoked potentials, an activity recorded at posterior electrodes co-occurs with a frontal activity of opposite sign if the equivalent dipoles have a posterior-anterior orientation (Joyce & Rossion, 2005). Thus, early visual activity can be characterized by a strong increase in the standard deviation across electrodes, a phenomenon captured by GFA.

### Statistical analyses

All statistical analyses were performed using a classic percentile bootstrap (Wilcox, 2005). First, we computed significant differences between two conditions averaged across subjects; second, differences between two conditions were estimated for each subject individually, using a single-trial approach.

Analyses across subjects were performed on the trimmed means by sampling subjects (i.e., sets of

observations from two conditions, consisting of the full electrodes by time points matrices) with replacement, averaging the trimmed means across subjects independently for each condition, and then computing the difference between the means for the two conditions (for instance upright faces vs. upright houses). This process was repeated 999 times, leading to a distribution of bootstrapped estimates of the mean difference between two ERP conditions, averaged across subjects. Then the 99% percent confidence interval was computed ( $\alpha = 0.01$ ). Finally, the difference between the two sample means was considered significant if the 99% confidence interval did not include zero. Note that this bootstrap technique, relying on an estimation of H1, tends to have more power than other robust methods like permutation tests and related bootstrap methods that evaluate the null hypothesis H0 (Wilcox, 2005).

Single-trial analyses were performed by sampling, independently with replacement, individual trials from the original distributions. Each sample consisted of the whole electrodes-by-time points matrix because, unlike trials, electrodes and time points are not independent from each other. Then the trimmed mean was computed for each condition and the difference between the two trimmed means stored. This process was repeated 999 times and a 99% confidence interval computed. Again, differences were considered significant if this interval did not contain zero.

Despite a systematic exploration of the data space, consisting of tracking differences at all electrodes and time points, no correction for multiple comparisons was applied. However, as will be evident in the result section, significant differences were not randomly scattered (as might be expected were these significant points the result of random type I error) but formed spatiotemporal clusters, as expected from our knowledge of the physics of EEG.

In addition to classic comparisons between sets of two categories, we estimated a measure of “single-trial reliability.” For each electrode and at each time point, we determined how many individual trials followed the pattern exhibited by the mean and the trimmed mean ERP. Concretely, if at a given time point (trimmed) mean ERP1 was superior to (trimmed) mean ERP2, then all the single trials from condition 1 received a score of 1 if they were superior to (trimmed) mean ERP2, otherwise they received a score of 0. Conversely, single trials from condition 2 received a score of 1 if they were inferior to (trimmed) mean ERP1, and 0 otherwise. Global reliability scores are then obtained at each electrode and each time point by averaging single-trial scores across trials and conditions. To provide a statistical evaluation of the reliability scores, the single trials from the two conditions were first pooled together and then sampled (400 sample trials) with replacement to determine the distribution of reliability scores obtained by chance under the null hypothesis according to which the two groups of trials

were sampled from the same population. From this distribution, a 99% confidence interval was obtained for each time point and electrode. Reliability scores that failed within the confidence interval for H0 did not contain any information about the difference observed between two ERP (trimmed) means and are masked in gray in all plots.

Finally, we used a measure of differential activity robustness that takes into account the single-trial variability observed in individual subjects. When comparing two conditions, instead of providing a simple binary output indicating whether an effect was or was not significant, we evaluated the probability of replicating such an effect. For each subject, this was achieved by sampling with replacement single trials from each condition before performing a bootstrap test ( $p < .01$ , 400 sample trials). The process of sampling single trials and applying a bootstrap test was performed 100 times (called Monte Carlo samples in Figure 8). In a sense, this procedure is equivalent to performing 100 fictive experiments. It follows from the logic according to which resampling the data in a bootstrap test is analog to the original data acquisition process. A measure of the robustness of differential activations was obtained for each subject individually and for group statistics. Because it was very time consuming, this analysis was carried out only on the GFA data in the first three experimental contrasts.

## Results

Subjects performed very well in this task: Mean accuracy was 95%, 96%, 95%, 95%, 96%, and 96% and median RT was 546 ms, 537 ms, 553 ms, 545 ms, 551 ms, and 556 ms, respectively, for inverted faces, inverted houses, inverted textures, upright faces, upright houses, and upright textures. No comparison reached significance. This is not surprising given that the task could be easily performed by discriminating simple elements such as a large oval for faces and a rectangle for houses. The simple discrimination task was used to maintain subjects' attention on the stimuli.

The differences among conditions were analyzed first using the GFA (the standard deviation across electrodes computed at each time point, see Methods). The GFA data for each separate condition are presented in Figure 2, in which three bursts of activity can be observed in the time window 100–300 ms, each corresponding in turn to the P1, N170, and P2 ERP components. The comparisons across all condition pairings are presented in Figure 3. GFA analyses revealed large periods of significant differences among conditions starting at 150 ms after stimulus onset. The first two rows of Figure 3 show small differences between object categories on the P1, the

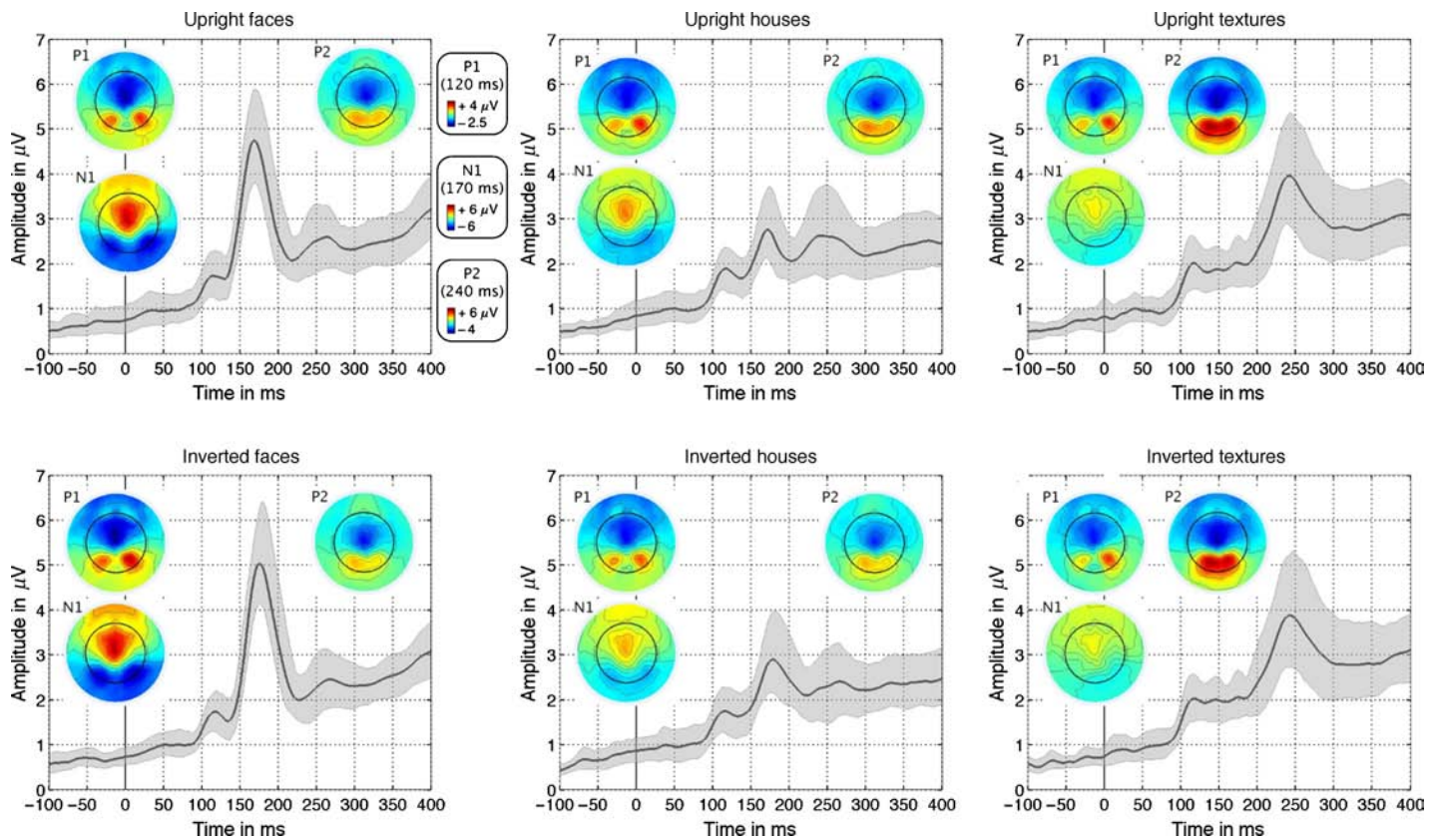


Figure 2. GFA for the 6 experimental conditions. Mean GFA and confidence interval of the mean are plotted with continuous lines and gray-shaded areas, respectively. Confidence intervals were computed using a percentile bootstrap with replacement, 1000 resampling trials, at  $p < .01$  (Wilcox, 2005). For each condition, the inserts show the mean topographic maps corresponding to the P1, N1, and P2 components.

successively decreased amplitude of the N170 from faces, to houses, and to textures, and finally, the much larger P2 for textures than for either faces or houses. In contrast to the category effects, inversion effects were rather subtle, as shown in the last row of Figure 3. In addition to the large differences in the N170 time window, some significant differences started as early as 100 ms after stimulus onset when faces (both upright and inverted) and houses (only inverted) were compared to textures.

Differences among conditions were analyzed in each subject individually, both on the GFA data and at each electrode. Figure 4 contains a summary of these analyses: for each pairwise comparison, Figure 4 shows the number of subjects who exhibited a significant difference. It is important to note that the GFA provides a very good description of the differences observed across electrodes, although its timing is conservative because it does not capture the onset of the earliest differences at individual electrodes. According to both measures, the most consistent differences were observed between 150 and 250 ms after stimulus onset, corresponding essentially to the time windows identified previously using the GFA across subjects. Between 140 and 200 ms, most subjects showed significant differences for faces compared to houses. When faces were compared to noise textures, those differences

started about 10 ms earlier (approximately 130 ms) and lasted up to 280 ms after stimulus onset. Face effects were found at most electrodes. In most subjects, differences between houses and textures emerged between 180 ms and 280 ms, or 30–50 ms after the onset of face–texture differences. These general patterns were present with both upright and inverted stimuli. In individual subjects, and contrary to the analyses performed across subjects, there was no evidence for significant effects before 130 ms. Also, the earliest differences between faces and houses occurred first over right hemisphere electrodes, a lateralization often reported in the literature (e.g., Jacques & Rossion, 2006). Inversion effects were much less consistent, both across subjects and across electrodes. They were centered on the N170 and on the P2 for faces, almost exclusively around the P2 for houses, and completely absent for noise textures.

Overall, our results indicate that the most reliable differences related to object processing occur in the time window of the N170. However, the measures we have used so far do not provide information about the robustness of the effects in individual subjects. Figure 5 reveals two important aspects of the ERP that differ substantially across subjects: the strength of the trimmed mean ERP differences between face, house, and texture conditions, and the width of the confidence interval around the trimmed mean for

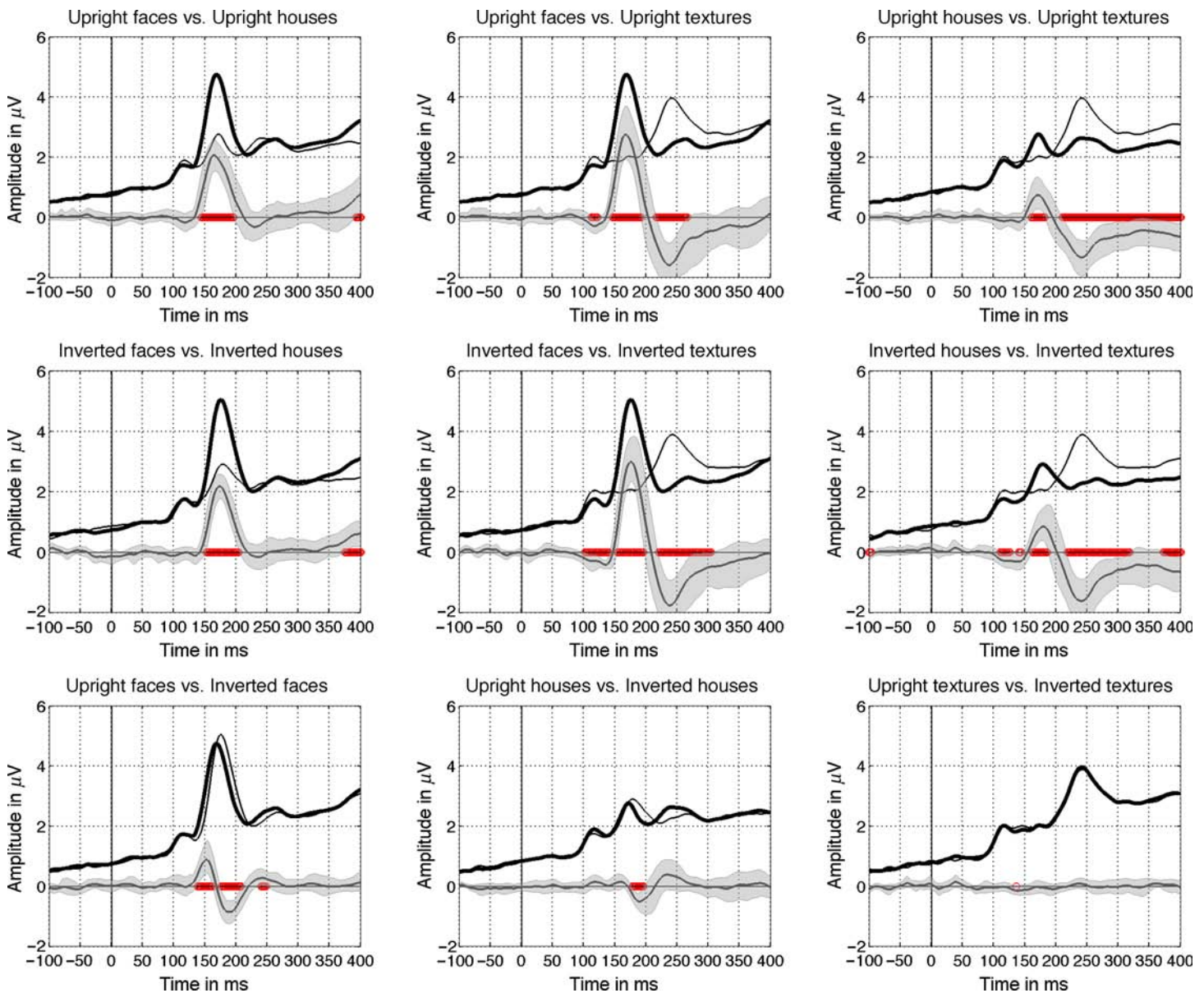


Figure 3. Comparisons of all condition pairings of GFA data. For each cell, the gray line is the difference between the conditions plotted in thick and thin black lines (respectively the first and the second element of the cell's title). The shaded gray area around the gray difference line is the confidence interval of the difference between the two conditions (percentile bootstrap, 1000 sample trials,  $p < .01$ ). When the confidence interval does not include zero, the difference is significant, as indicated by the thick horizontal red lines along the 0  $\mu V$ .

each condition. Because the trimmed means of two ERP conditions can be relatively close to each other, and with relatively large confidence intervals, we determined to what extent the difference observed between two trimmed means is actually present in single trials.

A measure of single-trial reliability of ERP differences is illustrated in Figures 6 and 7. First, we consider the single-trial reliability of the difference between upright faces and textures, which was the contrast yielding the largest effect. Subjects showed considerable variability in the size and the strength of clusters with a significant single-trial reliability. Despite those differences, all subjects showed a strong increase in reliability at 130–150 ms

after stimulus onset, consistent with the results observed with the GFA and ERP data. Figure 6 also demonstrates that trimmed means provide a strong increase in reliability compared to means. This is further illustrated in Figure 7, which presents results for all nine experimental contrasts at electrode P10. The single-trial reliability was stronger for the larger differential activities (faces compared to houses and textures), and for all subjects peaked in the time-window of the N170 or later. The reliability was particularly weak for the house/texture comparisons and the inversion effects. It is important to note that the increased single-trial reliability, in the baseline period, for trimmed means compared to means (Figure 7) does not

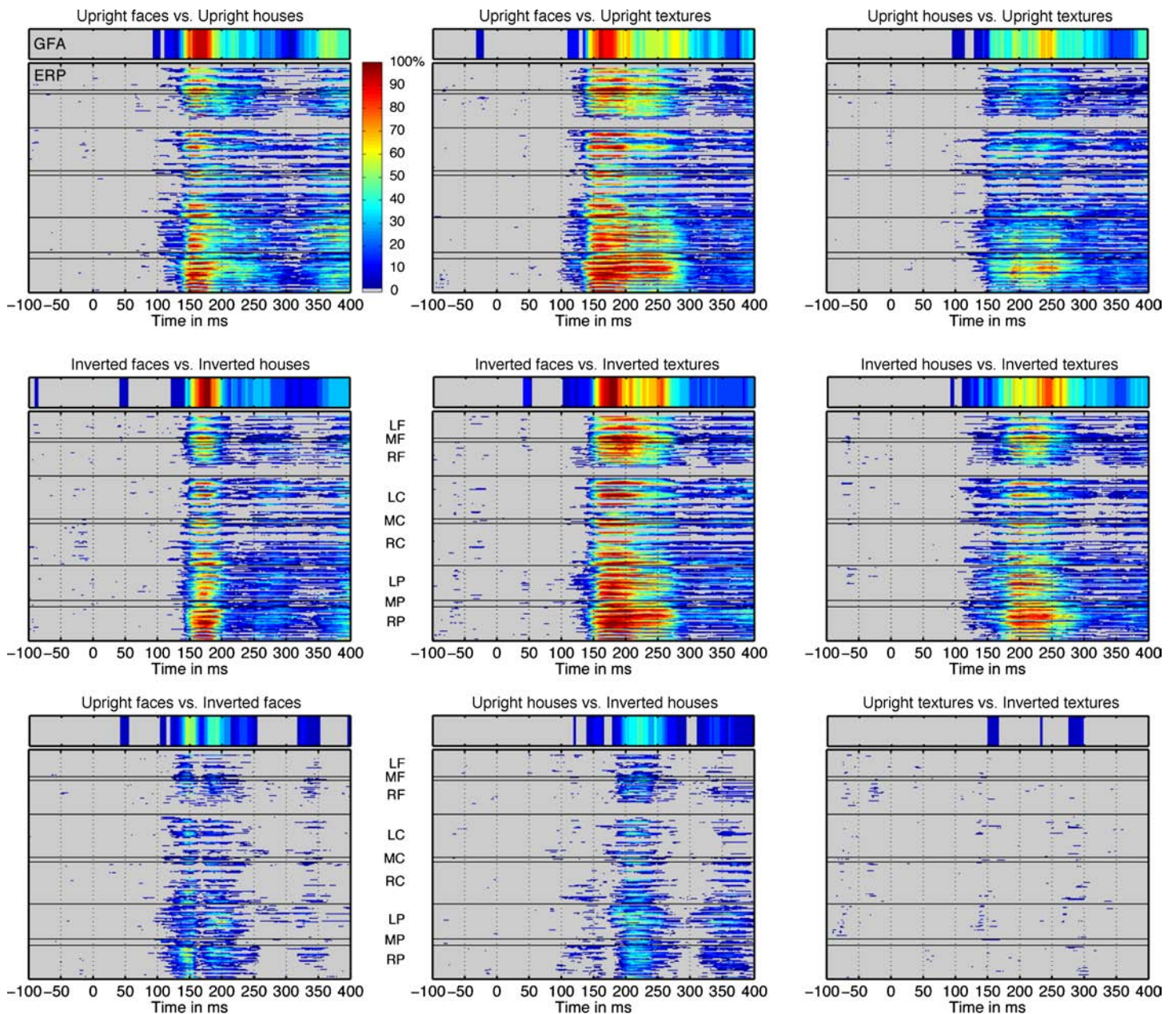


Figure 4. Number of subjects showing significant differences over time. The nine cells correspond to the same comparisons presented in Figure 3 (percentile bootstrap, 1000 sample trials,  $p < .01$ ). Each cell contains two subplots, one showing the GFA analyses (top), the other showing the ERP analyses at all electrodes. The color code is shown on the right side of the upper left cell. The number of subjects is coded from 0 to 100% because not all electrodes were available for each subject; i.e., it is the percentage of the subjects for whom a given electrode was available. Electrodes are stacked along the vertical axis. The horizontal black lines separate the different groups of electrodes organized in frontal, central, and posterior electrodes (F/C/P) and subdivided into left hemisphere, mid-line, and right hemisphere electrodes (L/M/R). Note that the patterns of differences follow our intuition that meaningful differences should be expressed across neighboring time points and electrodes.

reflect an increased sensitivity of trimmed-means to noise. By definition, because trimmed means average together a subset of trials that are more tightly clustered at the center of the original distribution, the single-trial variance of trimmed means is weaker than the variance for means. Thus, the patterns of reliability should not be confused with ERP differential activities, but should be evaluated in

conjunction with them. Overall, they clearly reinforce our previous observations showing that differences outside the time window of the N170 are inconsistent not only in individual subjects, but also in individual EEG trials.

We demonstrated previously that a simple model in which an evoked response of fixed amplitude is modulated by on-going activity could explain the single-trial variability



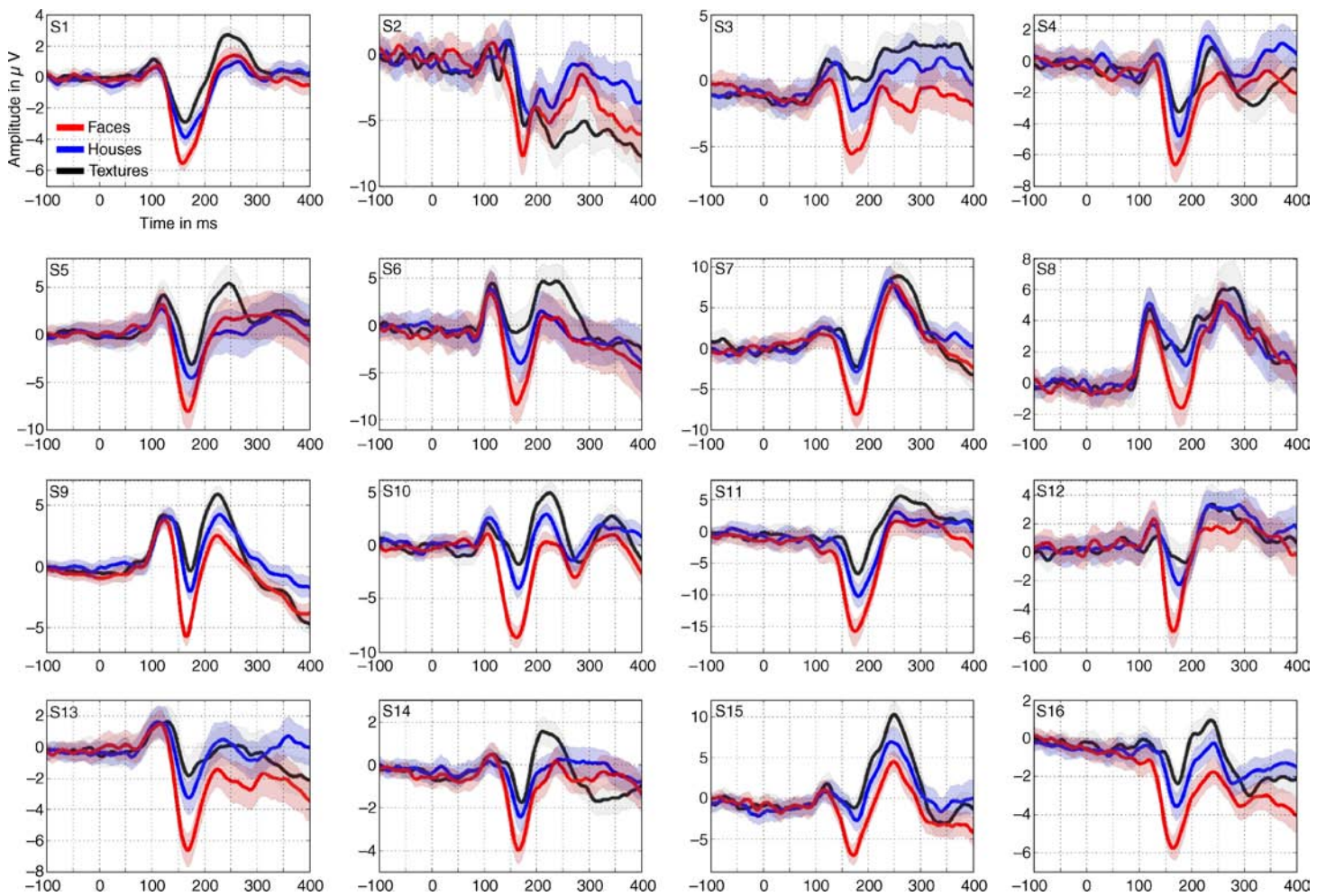
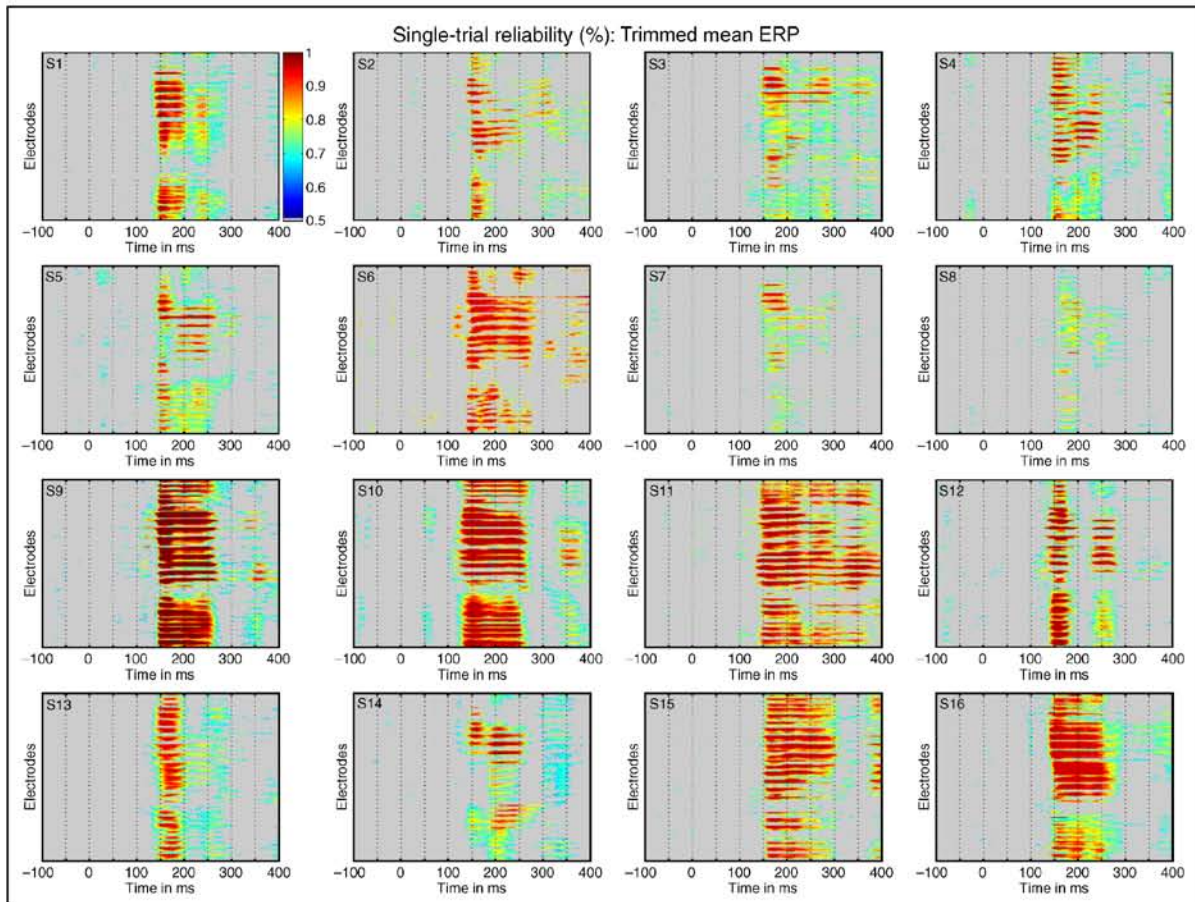
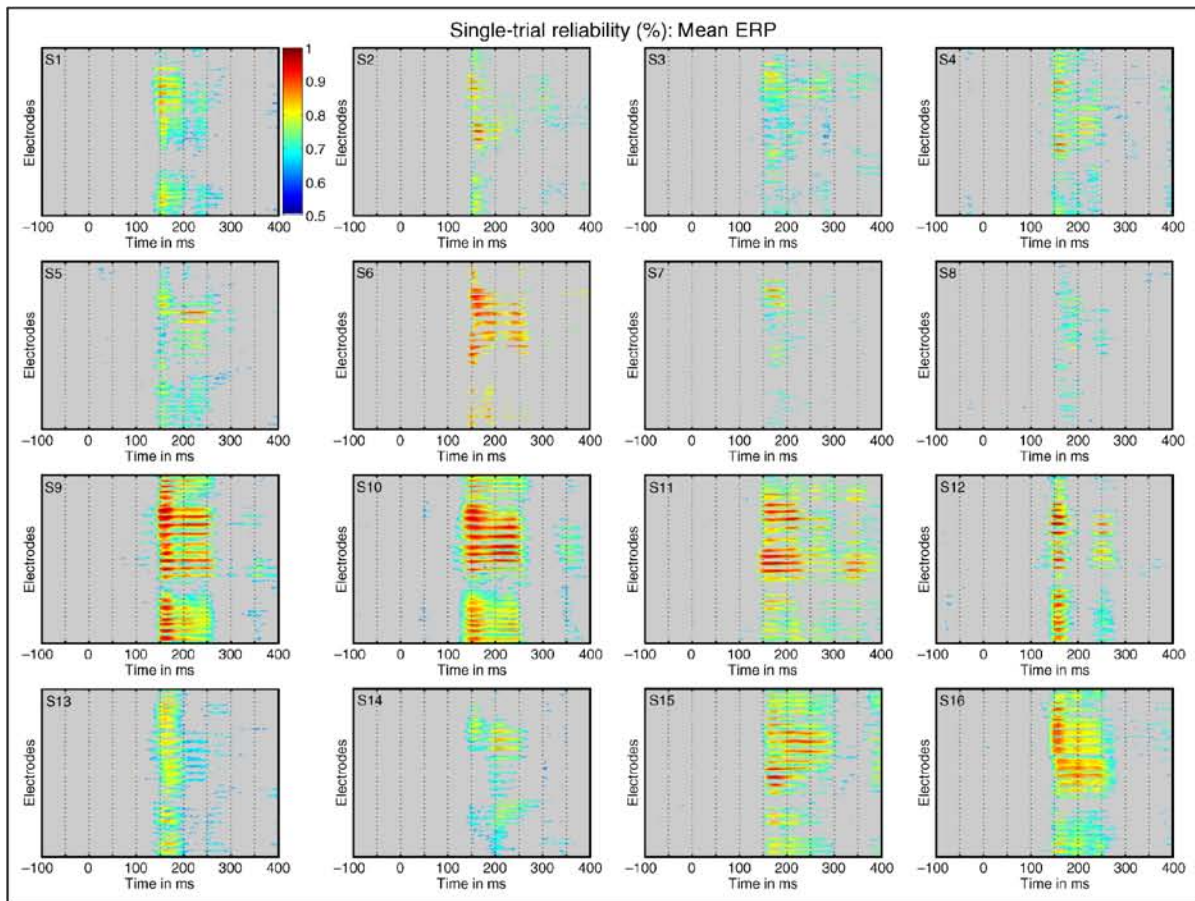


Figure 5. ERP data in individual subjects. Each cell shows one subject (S1–S16), with ERP to upright faces, houses, and textures represented in red, blue, and black, respectively. Thick lines represent the trimmed mean ERP, and shaded areas the 99% bootstrap confidence interval (percentile technique, 1000 sample trials). Data are from electrode E170 of the EGI system, corresponding to electrode P10, one of the right hemisphere electrodes presenting the strongest face effects across subjects.

observed in the present data and illustrated in [Figures 5](#), [Figures 6](#), [Figures 7](#) (Rousselet, Husk, et al., 2007). However, the statistical comparisons we have provided so far do not reflect this variability. Indeed, when performing a bootstrap test, for a chosen alpha level, at each time point the output is binary, indicating whether the difference is significant or not. Here we propose a strategy to measure significant differences between two ERP conditions that takes into account the single-trial variability ([Figure 8](#)). Instead of determining if two data points do or do not differ, our measure of differential activity robustness estimates the probability of finding a difference between those two points (see [Methods](#)). In [Figure 8](#), the color code ranges from dark blue to red corresponding to a range of probabilities from 0% to 100% of finding a significant difference between two conditions. This analysis is consistent with the results of the previous analyses but also revealed a striking discrepancy across subjects in terms of the timing, the duration, and the

robustness of the ERP differences. Importantly, statistics performed across subjects fail to capture this aspect of the data; they reflect only effects that are common to most subjects, masking potentially interesting patterns that are specific to a small proportion of the subjects. For instance, although the difference between faces and houses was statistically significant only in the range of 150–200 ms, several subjects clearly showed robust differences in the range of 200–300 ms. Regarding the comparison of houses and textures, it is also striking to see that the window of statistical significance 160–180 ms observed across subjects is almost completely absent in six subjects (S6, S7, S8, S12, S14, S15). Consider, also, the results shown in the bottom of the second column in [Figure 8](#): In each of the 100 Monte Carlo samples, differences between faces and textures were found consistently (i.e., in nearly every subject and nearly Monte Carlo sample) in two time windows around the N170 and the P2. On 45% of the Monte Carlo samples, a difference also was found in



the range of 110–130 ms after stimulus onset. However, unlike the differences near the N170 and P2, this earlier difference was found consistently in only two subjects (S4 and S10).

Finally, we measured the onset latency of the differential activities for the three contrasts reported in Figure 8 by defining the onset as the point at which more than 70% of significant Monte Carlo samples was obtained. The mean onsets with their 95% bootstrap confidence intervals were 144 ms [138–149] for faces vs. houses, 159 ms [141–187] for faces vs. textures and 188 ms [163–216] for houses vs. textures. The first two onset distributions did not differ from each other, but the onset for houses was significantly longer than those for faces ( $p < .01$ ).

## Discussion

### Time course of face and object processing

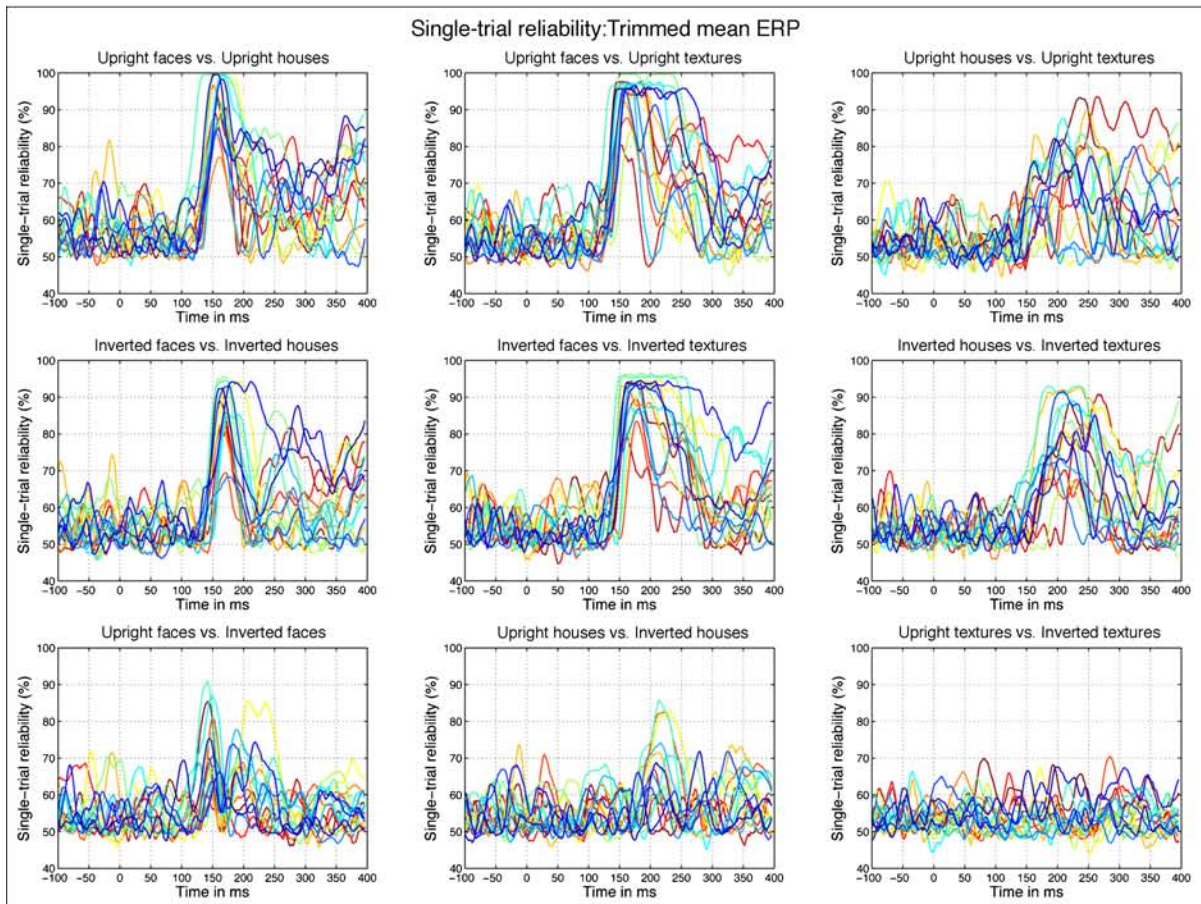
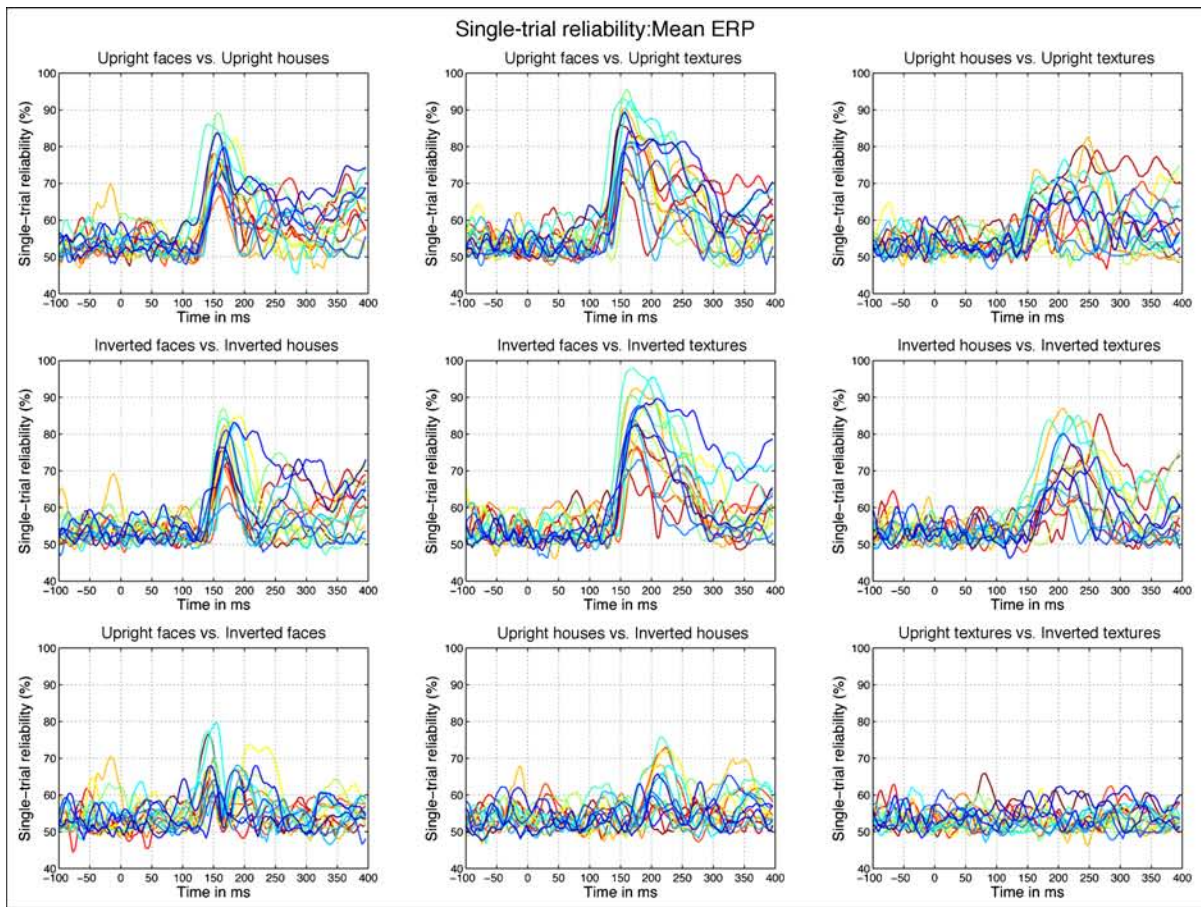
We investigated the time course of face and object processing using faces, houses, and noise textures with identical amplitude spectra. Our results confirm that the N170 face effect, as defined in terms of larger amplitude for faces compared to other objects, is not due to differences in amplitude spectra between faces and objects (Rousselet et al., 2005; Rousselet, Husk, et al. 2007). Here, this result is extended to the comparison between houses and noise, showing substantial ERP differences between houses and noise textures with identical amplitude spectra. We found significant differences between noise textures and faces starting at about 130–150 ms after stimulus onset. This timing is very similar to the ones reported in a number of previous studies using phase-scrambled noise (e.g., EEG: Jacques & Rossion, 2004,

Figure 6. Single-trial reliability of upright faces versus upright textures ERP differences. Data for the mean and the trimmed mean are presented in the top and bottom panels, respectively. Within each panel, each cell contains the data from one subject at all electrodes, with reliability scores averaged across trials and conditions. Non-significant reliability scores are masked in gray in all plots. This analysis demonstrates that the first reliable differences observed on the (trimmed) means start in the rising part of the N170, about 130–150 ms after stimulus onset. Importantly, trimming the data provides a strong increase in single-trial reliability, which is expected from a robust measure of location. This comparison is appropriate because it is only based on the capacity of the mean and the trimmed mean to capture the behavior of single trials, independently of their absolute values. However, a direct comparison between the two measures of location is impossible unless we know what the results ought to be. Such an absolute benchmark can be obtained using simulations and will be the topic of another paper.

2006, 2007; Philiastides & Sajda, 2006; MEG: Tanskanen et al., 2007). It is also consistent with previous EEG reports that used uncontrolled stimuli (Carmel & Bentin, 2002; Itier et al., 2006; Itier & Taylor, 2004; Rossion et al., 2000; Rousselet et al., 2004; Rousselet, Mace, et al., 2007; Schendan et al., 1998). Since we used gray-scale stimuli with submaximal contrast, the question remains open as to whether color and contrast might speed up processing speed even further (i.e., below 130 ms), as demonstrated in other studies (contrast: Macé, Thorpe, & Faber-Thope, 2005; color: Goffaux et al., 2003; see also monkey single-unit recording results in Edwards, Xiao, Keyser, Földiák, & Perrett, 2003). More generally, there is no reason to believe that amplitude spectra, colors, and textures cannot provide valuable information that can be used in certain classification tasks. But to the extent that faces and objects are still highly recognizable when presented in grayscale and equated in spectral energy, these stimulus constraints constitute an excellent way to tackle higher-order shape processing while eliminating a large range of low-level differences (Gold et al., 1999a, 1999b). In our experiment, there was no evidence whatsoever for significant differences before 130 ms. The only differences observed before that time point were weak in amplitude, scattered in space and time, and only present for at most a couple of observers.

We acknowledge that our conclusions are limited by the range of stimulus categories employed. It will be important to replicate the present finding with different object categories that vary in shape and familiarity. The comparison between faces and houses that we used has one strength and one limitation. First, because houses are known to produce a very different pattern of brain activity than faces (Haxby et al., 2001; Spiridon, Fischl, & Kanwisher, 2006), our analyses thus provide a potentially

Figure 7. Single-trial reliability of ERP differences for all contrasts. Single-trial reliability is shown for the mean (top half) and the trimmed mean (bottom half) at electrode E170 of the EGI geodesic system, corresponding to electrode P10 of the 10–20 system. In each cell, data from all subjects are shown in thin color lines. A thick black line represents the mean across subjects. Data were smoothed by a 5-steps running average for plotting purposes. The percentage in the upper left corner is the mean across subjects of the maximum single-trial reliability observed in the time window 100–300 ms. The 95% bootstrap confidence interval is in square brackets (percentile technique,  $\alpha = 0.05$ ). Compared to the mean, using the trimmed mean allowed a minimum gain in reliability of 5%, with a 95% confidence interval [3–6%] for the contrast upright compared to inverted textures (calculated using the maximum reliability in the time window 100–300 ms). A maximum gain of 16% [15–17%] was obtained for the upright faces compared to upright houses contrast. The two other contrasts for upright stimuli were associated with gains of 11% [9–13%] and 10% [8–11%] for faces compared to textures and houses compared to textures, respectively.



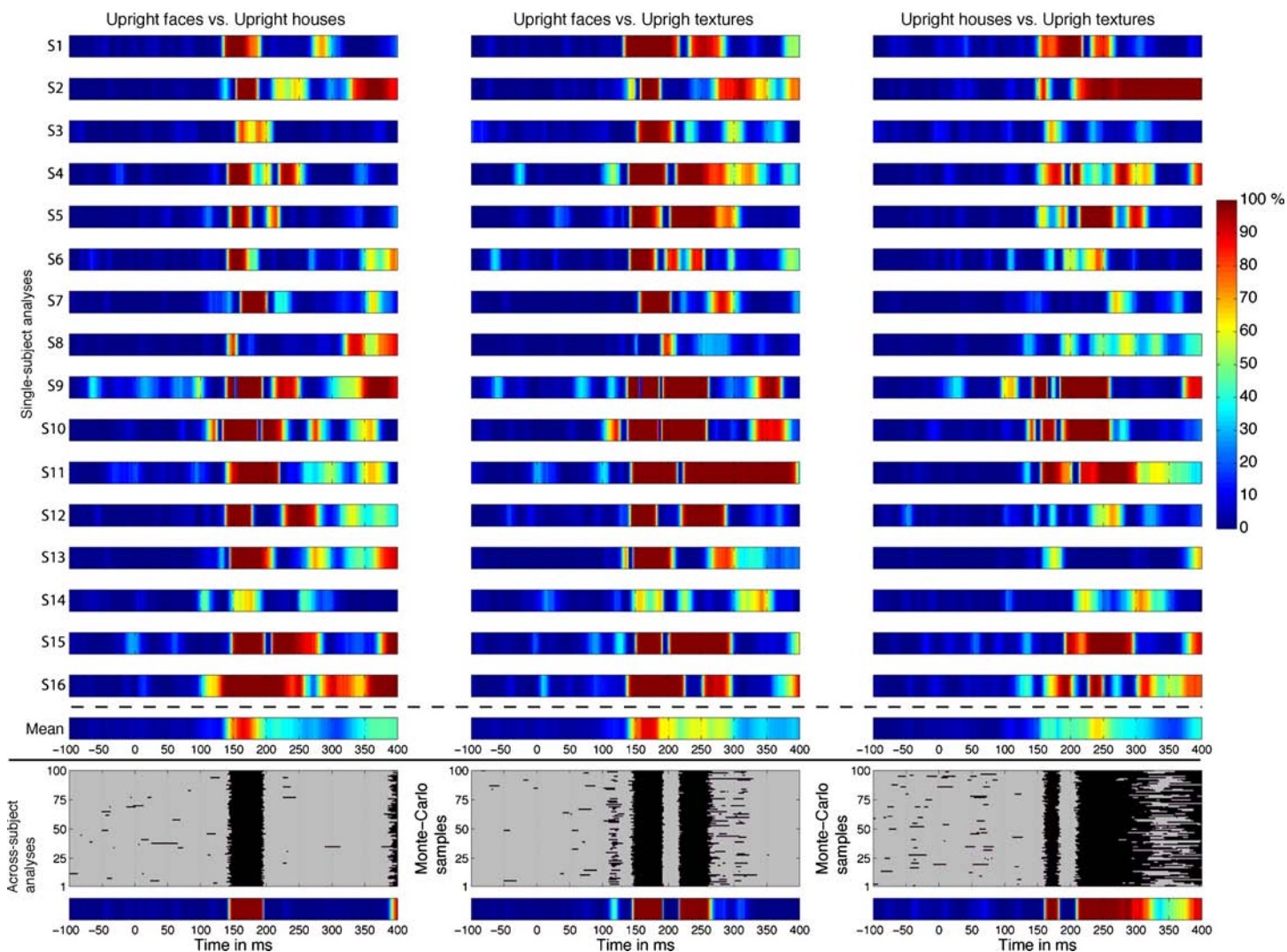


Figure 8. Robustness of ERP differential activities evaluated by a Monte Carlo simulation. Because a significant proportion of single trials do not show the effect observed on the trimmed means, it is somewhat misleading to make binary judgments about the statistical significance of an effect. This figure constitutes an alternative description of the data in terms of the probability of finding a difference at any time point between two conditions. The analysis was carried out on the GFA for each subject (S1–S16). The mean across the 16 subjects is plotted below the dashed line. The bottom of the figure depicts the result of the analysis performed across subjects. For each of the 100 samples in the simulation, the GFA for all subjects were used to compute an analysis across subjects, exactly like the one presented in Figure 3 ( $p < .01$ , 1000 sample trials). The three gray rectangles show, in black, the time points at which a significant difference between conditions, averaged across subjects, was observed for each of the 100 Monte Carlo samples. The color bars at the very bottom of the figure show the mean across the 100 Monte Carlo samples.

very liberal estimate of face processing speed. Second, other non-face objects might be more familiar and/or processed faster than houses. It is thus somewhat unfair to estimate non-face object processing speed from the house vs. texture comparison. Alternatively, one could also argue that using phase-randomized noise provides a very liberal estimate of processing speed because such noise lacks the local structure—created by edges, lines, and corners—and the long-range multi-scale correlations that exist in natural images (Field, 1999). However, if the differential activity evoked by faces and noise textures

was only related to the higher order statistical properties of natural images, then we would expect to find similar differences between houses and textures. The fact that differences between houses and textures were delayed by 30–50 ms relative to the differences between faces and textures speaks in favor of a real processing speed advantage for faces in the present experiment.

In terms of statistical analyses, we have implemented relatively simple steps that can be taken to insure the robustness of differential activities. These processing steps can be applied to differential activities produced by

differences in the stimulus and/or task. Our strategy was to address very simple questions starting with an exploration of differences observed across subjects, followed by an examination of the number of subjects showing a similar effect. It is worth noting that most EEG and MEG papers do not report the number of subjects showing an effect, although there is a recent trend in that direction (e.g., Philiastides & Sajda, 2006; Schyns, Petro, & Smith, 2007; Smith, Gosselin, & Schyns, 2007). Ensuring that an effect is observed across all or most subjects is essential because in some situations (for instance the early P1 difference we report in the present paper) an effect might be driven by a minority of subjects. This is not to say that such effects are not interesting, but rather that they should be interpreted with caution. At the single-subject level, we have described two strategies to ensure the strength of differential activities. The first strategy, single-trial reliability (Figure 7), relies on an estimation of the number of trials following the same pattern observed at the level of the measure of central tendency. Usually, the mean is used as a measure of central tendency. Here we demonstrate that using other measures of central tendency, such as the trimmed mean, might provide a better description of the data, following the work of Wilcox (2005). The second strategy, single-trial robustness (Figure 8), relies on Monte Carlo simulations to determine the probability of observing an effect. This kind of probabilistic description of the results might seem unusual but it is a much more faithful description of the data, given that, strictly speaking, statistical analyses based upon probability distributions can only be used to estimate our chances to replicate a given result if we were to repeat the experiment, not to validate or falsify an hypothesis (Goodman, 1999). Finally, we also note the existence of complementary strategies, for instance relying on two-level hierarchical models that estimate both the within-subject variability and the inter-subject variability to provide group statistics (Kiebel & Friston, 2004a, 2004b).

Overall, even if it is difficult to draw conclusions from null effects, our range of analyses, as well as congruent reports from different groups, lead us to conclude that there is no reason to believe that face-specific responses occur before about 130 ms after stimulus onset. In addition, it is important to remember that onsets reported here and in other papers correspond to time points at which significant differences are first observed. Differential activities generally keep increasing for several tens of milliseconds, which might correspond to a process of accumulation of information.

## Inversion effect

Evidence for the onset of face-sensitive processing can also be obtained from the inversion effect. Although inversion effects for objects other than faces have been reported previously (Eimer, 2000a; Itier et al., 2006;

Rossion et al., 2003; Rousselet et al., 2004), inversion effects of large amplitude seem to be a hallmark of human faces (Itier et al., 2006; Rousselet et al., 2004). Such face-sensitive inversion effects have been reported as early as P1 (Itier & Taylor, 2002, 2004; Linkenkaer-Hansen et al., 1998), though most studies report such differences nearer to the N170 (Itier et al., 2006; Jacques & Rossion, 2007; Rossion & Gauthier, 2002; Rousselet et al., 2004). Here, we found that the N170 had a longer latency and larger amplitude for inverted face than upright faces, but that face inversion had minimal effects on earlier parts of the ERP. We also observed an inversion effect for houses, but it occurred in the time window of the P2, not the N170, and was smaller than the face inversion effect. Textures, as expected, did not produce an inversion effect, as they did not have an a priori canonical orientation. Thus, the effects of stimulus inversion were consistent with the hypothesis that face-preferential aspects of the ERP emerge within the time window of the N170.

## Later activity

Following the N170, the P2 ERP component was modulated strongly by stimulus category: P2 amplitude was consistently larger in response to noise textures than to both faces and houses, irrespective of orientation. The literature generally refers to the P2 as reflecting later stages of face processing compared to the N170, such as recognition or decision in the context of the task (Halit et al., 2000; Itier & Taylor, 2002; Latinus & Taylor, 2005). Alternatively, the P2 might reflect, in part, the activation of the same cortical patches generating the N170, as suggested by Itier, Herdman, George, Cheyne, and Taylor (2006). Hence, the larger P2 for textures might reflect the same mechanisms underlying the N170 for faces and houses but delayed in time. The hypothesis also is supported by a recent study showing, in monkeys, strong local field potential responses to faces peaking at about 130 and 200 ms after stimulus onset (Tsao, Freiwald, Tootell, & Livingstone, 2006). Since surface EEG measures local field potentials (Shah et al., 2004), it is tempting to link those two peaks to the N170 and P2 components, recorded in our subjects at about 160–180 ms and 230–250 ms after stimulus onset. The delay between activation in monkeys and in humans could be due to the larger size of human brains (Fabre-Thorpe, Richard, & Thorpe, 1998).

## Conclusion

In sum, assessing ERPs produced by different stimulus categories in each subject individually revealed that consistent differential effects—i.e., ERP differences that were observed in almost all subjects—occurred only in the

N170 range, starting at about 130–150 ms for faces vs. houses and about 30–40 ms later for houses vs. textures. Here, we do not intend to exclude the possibility of “true” object related differences before the N170. However, such differences will have to be analyzed systematically in individual subjects to evaluate their reliability and by varying image parameters to determine to what physical properties they might correspond. The simple measures of single-trial reliability we have introduced in this paper constitute important criteria to define “interesting” evoked responses because one should be able eventually to read out the information about object categories from single trials, and not only from mean evoked activity (Schyns, Jentzsch, Johnson, Schweinberger, & Gosselin, 2003; Piliastides & Sajda, 2006). After all, the brain does its job on each single trial and does not accumulate information across the whole experiment before providing a response. Finally, our paper did not touch the topic of the specificity or the nature of the brain responses described. Our goal was to provide a robust time line for the extraction of some coarsely defined object information. By working in a parametrically controlled task and stimulus information space one can get at the underlying mechanisms, as well as their categorical specificity (Pernet, Schyns, & Demonet, 2007). The rather crude categorical comparisons reported here and in the vast majority of the literature published so far fall short in this respect.

## Acknowledgments

This work was supported by NSERC Discovery Grants 42133 and 105494, Canada Research Chairs, and infrastructure support from the CFI and OIT to PJB and ABS, an NSERC PGS-D award to JSH, and by a CIHR fellowship grant to GAR. The McMaster Research Ethics Board approved this research.

Commercial relationships: none.

Corresponding author: Guillaume A. Rousselet.

Email: g.rousselet@psy.gla.ac.uk.

Address: 58 Hillhead Street, G12 9XW, Glasgow, UK.

## References

Allison, T., Puce, A., Spencer, D. D., & McCarthy, G. (1999). Electrophysiological studies of human face perception: I. Potentials generated in occipitotemporal cortex by face and non-face stimuli. *Cerebral Cortex*, *9*, 415–430. [[PubMed](#)] [[Article](#)]

Bentin, S., & Golland, Y. (2002). Meaningful processing of meaningless stimuli: The influence of perceptual experience on early visual processing of faces. *Cognition*, *86*, B1–B14. [[PubMed](#)]

Bötzel, K., Schulze, S., & Stodieck, S. R. (1995). Scalp topography and analysis of intracranial sources of face-evoked potentials. *Experimental Brain Research*, *104*, 135–143. [[PubMed](#)]

Carmel, D., & Bentin, S. (2002). Domain specificity versus expertise: Factors influencing distinct processing of faces. *Cognition*, *83*, 1–29. [[PubMed](#)]

Cauquil, A. S., Edmonds, G. E., & Taylor, M. J. (2000). Is the face-sensitive N170 the only ERP not affected by selective attention? *Neuroreport*, *11*, 2167–2171. [[PubMed](#)]

Debruille, J. B., Guillem, F., & Renault, B. (1998). ERPs and chronometry of face recognition: Following-up Seeck et al. and George et al. *Neuroreport*, *9*, 3349–3353. [[PubMed](#)]

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*, 9–21. [[PubMed](#)]

Edwards, R., Xiao, D., Keysers, C., Földiák, P., & Perrett, D. (2003). Color sensitivity of cells responsive to complex stimuli in the temporal cortex. *Journal of Neurophysiology*, *90*, 1245–1256. [[PubMed](#)] [[Article](#)]

Eimer, M. (2000a). Effects of face inversion on the structural encoding and recognition of faces. Evidence from event-related brain potentials. *Brain Research: Cognitive Brain Research*, *10*, 145–158. [[PubMed](#)]

Eimer, M. (2000b). Attentional modulations of event-related brain potentials sensitive to faces. *Cognitive Neuropsychology*, *17*, 103–116.

Fabre-Thorpe, M., Richard, G., & Thorpe, S. J. (1998). Rapid categorization of natural images by rhesus monkeys. *Neuroreport*, *9*, 303–308. [[PubMed](#)]

Field, D. (1999). Wavelets, vision and the statistics of natural scenes. *Philosophical Transactions of the Royal Society of London A*, *357*, 2527–2542.

Furey, M. L., Tanskanen, T., Beauchamp, M. S., Avikainen, S., Uutela, K., Hari, R., et al. (2006). Dissociation of face-selective cortical responses by attention. *Proceedings of the National Academy of Sciences of the United States of America*, *103*, 1065–1070. [[PubMed](#)] [[Article](#)]

George, N., Jemel, B., Fiori, N., & Renault, B. (1997). Face and shape repetition effects in humans: A spatio-temporal ERP study. *Neuroreport*, *8*, 1417–1423. [[PubMed](#)]

Goffaux, V., Gauthier, I., & Rossion, B. (2003). Spatial scale contribution to early visual differences between face and object processing. *Brain Research: Cognitive Brain Research*, *16*, 416–424. [[PubMed](#)]

Gold, J., Bennett, P. J., & Sekuler, A. B. (1999a). Identification of band-pass filtered letters and faces

- by human and ideal observers. *Vision Research*, *39*, 3537–3560. [PubMed]
- Gold, J., Bennett, P. J., & Sekuler, A. B. (1999b). Signal but not noise changes with perceptual learning. *Nature*, *402*, 176–178. [PubMed]
- Goodman, S. N. (1999). Toward evidence-based medical statistics: I. The P value fallacy. *Annals of Internal Medicine*, *130*, 995–1004. [PubMed]
- Halit, H., de Haan, M., & Johnson, M. H. (2000). Modulation of event-related potentials by prototypical and atypical faces. *Neuroreport*, *11*, 1871–1875. [PubMed]
- 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*, 2425–2430. [PubMed]
- 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). *Journal of Neural Transmission*, *112*, 1073–1081. [PubMed]
- Husk, J. S., Bennett, P. J., & Sekuler, A. B. (2007). Inverting houses and textures: Investigating the characteristics of learned inversion effects. *Vision Research*, *47*, 3350–3359. [PubMed]
- Itier, R. J., Herdman, A. T., George, N., Cheyne, D., & Taylor, M. J. (2006). Inversion and contrast-reversal effects on face processing assessed by MEG. *Brain Research*, *1115*, 108–120. [PubMed]
- Itier, R. J., Latinus, M., & Taylor, M. J. (2006). Face, eye and object early processing: What is the face specificity? *Neuroimage*, *29*, 667–676. [PubMed]
- Itier, R. J., & Taylor, M. J. (2002). Inversion and contrast polarity reversal affect both encoding and recognition processes of unfamiliar faces: A repetition study using ERPs. *Neuroimage*, *15*, 353–372. [PubMed]
- Itier, R. J., & Taylor, M. J. (2004). N170 or N1? Spatiotemporal differences between object and face processing using ERPs. *Cerebral Cortex*, *14*, 132–142. [PubMed] [Article]
- Jacques, C., & Rossion, B. (2004). Concurrent processing reveals competition between visual representations of faces. *Neuroreport*, *15*, 2417–2421. [PubMed]
- Jacques, C., & Rossion, B. (2006). The time course of visual competition to the presentation of centrally fixated faces. *Journal of Vision*, *6*(2):6, 154–162, <http://journalofvision.org/6/2/6/>, doi:10.1167/6.2.6. [PubMed] [Article]
- Jacques, C., & Rossion, B. (2007). Early electrophysiological responses to multiple face orientations correlate with individual discrimination performance in humans. *Neuroimage*, *36*, 863–876. [PubMed]
- Jeffreys, D. (1996). Evoked potential studies of face and object processing. *Visual Cognition*, *3*, 1–38.
- Johnson, J. S., & Olshausen, B. A. (2003). Timecourse of neural signatures of object recognition. *Journal of Vision*, *3*(7):4, 499–512, <http://journalofvision.org/3/7/4/>, doi:10.1167/3.7.4. [PubMed] [Article]
- Joyce, C., & Rossion, B. (2005). The face-sensitive N170 and VPP components manifest the same brain processes: The effect of reference electrode site. *Clinical Neurophysiology*, *116*, 2613–2631. [PubMed]
- Kiani, R., Esteky, H., & Tanaka, K. (2005). Differences in onset latency of macaque inferotemporal neural responses to primate and non-primate faces. *Journal of Neurophysiology*, *94*, 1587–1596. [PubMed] [Article]
- Kiebel, S. J., & Friston, K. J. (2004a). Statistical parametric mapping for event-related potentials: I. Generic considerations. *Neuroimage*, *22*, 492–502. [PubMed]
- Kiebel, S. J., & Friston, K. J. (2004b). Statistical parametric mapping for event-related potentials (II): A hierarchical temporal model. *Neuroimage*, *22*, 503–520. [PubMed]
- Latinus, M., & Taylor, M. J. (2005). Holistic processing of faces: Learning effects with Mooney faces. *Journal of Cognitive Neuroscience*, *17*, 1316–1327. [PubMed]
- Lehmann, D., & Skrandies, W. (1980). Reference-free identification of components of checkerboard-evoked multichannel potential fields. *Electroencephalography and Clinical Neurophysiology*, *48*, 609–621. [PubMed]
- 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. *Neuroscience Letters*, *253*, 147–150. [PubMed]
- Logothetis, N. K., & Sheinberg, D. L. (1996). Visual object recognition. *Annual Review of Neuroscience*, *19*, 577–621. [PubMed]
- Luck, S. J., Woodman, G. F., & Vogel, E. K. (2000). Event-related potential studies of attention. *Trends in Cognitive Sciences*, *4*, 432–440. [PubMed]
- Lueschow, A., Sander, T., Boehm, S. G., Nolte, G., Trahms, L., & Curio, G. (2004). Looking for faces: Attention modulates early occipitotemporal object processing. *Psychophysiology*, *41*, 350–360. [PubMed]
- Macé, M. J., Thorpe, S. J., & Fabre-Thorpe, M. (2005). Rapid categorization of achromatic natural scenes: How robust at very low contrasts? *European Journal of Neuroscience*, *21*, 2007–2018. [PubMed]



- Mouchetant-Rostaing, Y., Giard, M. H., Bentin, S., Aguera, P. E., & Pernier, J. (2000). Neurophysiological correlates of face gender processing in humans. *European Journal of Neuroscience*, *12*, 303–310. [PubMed]
- Mouchetant-Rostaing, Y., Giard, M. H., Delpuech, C., Echallier, J. F., & Pernier, J. (2000). Early signs of visual categorization for biological and non-biological stimuli in humans. *Neuroreport*, *11*, 2521–2525. [PubMed]
- Nobre, A. C., Allison, T., & McCarthy, G. (1998). Modulation of human extrastriate visual processing by selective attention to colours and words. *Brain*, *121*, 1357–1368. [PubMed] [Article]
- Oppenheim, A. V., & Lim, J. S. (1981). The importance of phase in signals. *Proceedings of the IEEE*, *69*, 529–541.
- Pernet, C., Schyns, P. G., & Demonet, J. F. (2007). Specific, selective or preferential: Comments on category specificity in neuroimaging. *Neuroimage*, *35*, 991–997. [PubMed]
- Philiastides, M. G., & Sajda, P. (2006). Temporal characterization of the neural correlates of perceptual decision making in the human brain. *Cerebral Cortex*, *16*, 509–518. [PubMed] [Article]
- Puce, A., Allison, T., & McCarthy, G. (1999). Electrophysiological studies of human face perception: III. Effects of top-down processing on face-specific potentials. *Cerebral Cortex*, *9*, 445–458. [PubMed] [Article]
- Rossion, B., & Gauthier, I. (2002). How does the brain process upright and inverted faces? *Behavioral and Cognitive Neuroscience Reviews*, *1*, 63–75. [PubMed]
- 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*, 69–74. [PubMed]
- 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*, 1609–1624. [PubMed]
- Rousselet, G. A., Macé, M. J., & Fabre-Thorpe, M. (2004). Animal and human faces in natural scenes: How specific to human faces is the N170 ERP component? *Journal of Vision*, *4*(1):2, 13–21, <http://journalofvision.org/4/1/2/>, doi:10.1167/4.1.2. [PubMed] [Article]
- Rousselet, G. A., Husk, J. S., Bennett, P. J., & Sekuler, A. B. (2005). Spatial scaling factors explain eccentricity effects on face ERPs. *Journal of Vision*, *5*(10):1, 755–763, <http://journalofvision.org/5/10/1/>, doi:10.1167/5.10.1. [PubMed] [Article]
- Rousselet, G. A., Husk, J. S., Bennett, P. J., & Sekuler, A. B. (2007). Single-trial EEG dynamics of object and face visual processing. *Neuroimage*, *36*, 843–862. [PubMed]
- Rousselet, G. A., Macé, M. J., Thorpe, S. J., & Fabre-Thorpe, M. (2007). Limits of event-related potential differences in tracking object processing speed. *Journal of Cognitive Neuroscience*, *19*, 1241–1258. [PubMed]
- Schendan, H. E., Ganis, G., & Kutas, M. (1998). Neurophysiological evidence for visual perceptual categorization of words and faces within 150 ms. *Psychophysiology*, *35*, 240–251. [PubMed]
- Schyns, P. G., Jentzsch, I., Johnson, M., Schweinberger, S. R., & Gosselin, F. (2003). A principled method for determining the functionality of brain responses. *Neuroreport*, *14*, 1665–1669. [PubMed]
- Schyns, P. G., Petro, L. S., & Smith, M. L. (2007). Dynamics of visual information integration in the brain for categorizing facial expressions. *Current Biology*, *17*, 1580–1585. [PubMed]
- 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*, 2749–2754. [PubMed]
- Sekuler, A. B., & Bennett, P. J. (1996). Spatial phase differences can drive apparent motion. *Perception & Psychophysics*, *58*, 174–190. [PubMed]
- Shah, A. S., Bressler, S. L., Knuth, K. H., Ding, M., Mehta, A. D., Ulbert, I., et al. (2004). Neural dynamics and the fundamental mechanisms of event-related brain potentials. *Cerebral Cortex*, *14*, 476–483. [PubMed] [Article]
- Smith, M. L., Gosselin, F., & Schyns, P. G. (2007). From a face to its category via a few information processing states in the brain. *Neuroimage*, *37*, 974–984. [PubMed]
- Spiridon, M., Fischl, B., & Kanwisher, N. (2006). Location and spatial profile of category-specific regions in human extrastriate cortex. *Human Brain Mapping*, *27*, 77–89. [PubMed]
- Tanskanen, T., Näsänen, R., Montez, T., Päällysaho, J., & Hari, R. (2005). Face recognition and cortical responses show similar sensitivity to noise spatial frequency. *Cerebral Cortex*, *15*, 526–534. [PubMed] [Article]
- Tanskanen, T., Näsänen, R., Ojanpää, H., & Hari, R. (2007). Face recognition and cortical responses: Effect of stimulus duration. *Neuroimage*, *35*, 1636–1644. [PubMed]

- Tsao, D. Y., Freiwald, W. A., Tootell, R. B., & Livingstone, M. S. (2006). A cortical region consisting entirely of face-selective cells. *Science*, *311*, 670–674. [[PubMed](#)]
- Tucker, D. M. (1993). Spatial sampling of head electrical fields: The geodesic sensor net. *Electroencephalography and Clinical Neurophysiology*, *87*, 154–163. [[PubMed](#)]
- VanRullen, R., & Thorpe, S. J. (2001). The time course of visual processing: From early perception to decision-making. *Journal of Cognitive Neuroscience*, *13*, 454–461. [[PubMed](#)]
- Wilcoxon, R. R. (2005). *Introduction to Robust Estimation and Hypothesis Testing* (2nd Ed.). San Diego: Academic Press.
- Wilcoxon, R. R., & Keselman, H. J. (2003). Modern robust data analysis methods: Measures of central tendency. *Psychological Methods*, *8*, 254–274. [[PubMed](#)]