The N170 is Sensitive to Long-term (Personal) Familiarity of a Face Identity

Stéphanie Caharel a,b and Bruno Rossion c,d*

a Université de Lorraine, Laboratoire Lorrain de Psychologie et Neurosciences (2LPN - EA 7489), Nancy F-54000, France
b Institut Universitaire de France, Paris F-75000, France
c Université de Lorraine, CNRS, CRAN, F-54000 Nancy, France
d CHRU-Nancy, Service de Neurologie, F-54000, France

Abstract—The N170 is a large deflection of the human electroencephalogram (EEG), peaking at about 170 milliseconds over the occipito-temporal cortex after the sudden onset of a face stimulus. The N170 reflects perceptual awareness of a face and its onset corresponds to the emergence of reliable face-selectivity in the human brain. However, whether sensitivity to the long-term familiarity of a face identity emerges already at this early time-point remains debated. Here we provide a brief survey of the 45 published studies comparing the N170 response to unfamiliar and familiar (famous, experimentally familiarized, personally familiar and own) faces. Even though effects of familiarity on the N170 are relatively small and inconsistent across studies, this overview indicates that face familiarity significantly increases the N170 amplitude. This effect is especially present for personally familiar faces, learned in natural conditions. In the human brain, effects linked to familiarity with specific facial identities therefore appear to emerge between 150 and 200 ms in occipito-temporal brain regions, i.e., shortly after the onset of face-selectivity but at the same time as the earliest high-level effects of immediate unfamiliar face identity repetition. This observation challenges standard neurocognitive models with a clear-cut distinction between perceptual and memory stages in human face recognition. © 2021 IBRO. Published by Elsevier Ltd. All rights reserved.

Key words: event-related potential, N170, face familiarity, face identity, EEG.

INTRODUCTION

The N170 event-related potential (ERP) component is a relatively large deflection of the human electroencephalogram (EEG) peaking on average at about 170 milliseconds (ms) after the sudden onset of a face stimulus (with an interindividual variability range between 130 and 200 ms). Initially described by Srebro (1985) in an elegant study linking this ERP to face perception, i.e. the subjective experience of seeing a stimulus as a face, it was defined later on by Bentin and colleagues (Bentin et al., 1996), who established the larger N170 amplitude to faces than other familiar visual shapes (Fig. 1; see Rossion and Jacques, 2008 for a review of the face-selectivity of the N170). The N170 is maximal in amplitude on recording electrodes positioned over the occipito-temporal cortex, with generally a larger response in the right hemisphere, and emerges concurrently to a positive potential recorded on central electrode sites, the vertex positive potential (VPP; Botzel and Grusser, 1989; Jeffreys, 1989; see Jeffreys, 1996 for review). The N170 and VPP typically show similar response properties (Joyce and Rossion, 2005). One of the most salient, category-selective, response properties of the N170 is its larger amplitude and delayed latency following picture-plane inversion of a face stimulus (e.g., Rossion et al., 2000; Itier and Taylor, 2002). Over the last four decades, the N170 component has been described in hundreds, perhaps more than a thousand, of studies (Eimer, 2011; Rossion and Jacques, 2011 for reviews). In recent years, the N170 has been used as an index of deficient face recognition abilities in clinical populations, with moderate limited success (e.g., in Autism Spectrum Disorders (ASD): Batty et al., 2011; Jeffreys, 1989; see Jeffreys, 1996 for review). The N170 and VPP typically show similar response properties (Joyce and Rossion, 2005). One of the most salient, category-selective, response properties of the N170 is its larger amplitude and delayed latency following picture-plane inversion of a face stimulus (e.g., Rossion et al., 2000; Itier and Taylor, 2002). Over the last four decades, the N170 component has been described in hundreds, perhaps more than a thousand, of studies (Eimer, 2011; Rossion and Jacques, 2011 for reviews). In recent years, the N170 has been used as an index of deficient face recognition abilities in clinical populations, with moderate limited success (e.g., in Autism Spectrum Disorders (ASD): Batty et al., 2011; Jeffreys, 1989; see Jeffreys, 1996 for review). The N170 and VPP typically show similar response properties (Joyce and Rossion, 2005). One of the most salient, category-selective, response properties of the N170 is its larger amplitude and delayed latency following picture-plane inversion of a face stimulus (e.g., Rossion et al., 2000; Itier and Taylor, 2002). Over the last four decades, the N170 component has been described in hundreds, perhaps more than a thousand, of studies (Eimer, 2011; Rossion and Jacques, 2011 for reviews). In recent years, the N170 has been used as an index of deficient face recognition abilities in clinical populations, with moderate limited success (e.g., in Autism Spectrum Disorders (ASD): Batty et al., 2011; Jeffreys, 1989; see Jeffreys, 1996 for review). The N170 and VPP typically show similar response properties (Joyce and Rossion, 2005). One of the most salient, category-selective, response properties of the N170 is its larger amplitude and delayed latency following picture-plane inversion of a face stimulus (e.g., Rossion et al., 2000; Itier and Taylor, 2002). Over the last four decades, the N170 component has been described in hundreds, perhaps more than a thousand, of studies (Eimer, 2011; Rossion and Jacques, 2011 for reviews). In recent years, the N170 has been used as an index of deficient face recognition abilities in clinical populations, with moderate limited success (e.g., in Autism Spectrum Disorders (ASD): Batty et al., 2011; Jeffreys, 1989; see Jeffreys, 1996 for review). The N170 and VPP typically show similar response properties (Joyce and Rossion, 2005). One of the most salient, category-selective, response properties of the N170 is its larger amplitude and delayed latency following picture-plane inversion of a face stimulus (e.g., Rossion et al., 2000; Itier and Taylor, 2002). Over the last four decades, the N170 component has been described in hundreds, perhaps more than a thousand, of studies (Eimer, 2011; Rossion and Jacques, 2011 for reviews). In recent years, the N170 has been used as an index of deficient face recognition abilities in clinical populations, with moderate limited success (e.g., in Autism Spectrum Disorders (ASD): Batty et al., 2011; Jeffreys, 1989; see Jeffreys, 1996 for review). The N170 and VPP typically show similar response properties (Joyce and Rossion, 2005). One of the most salient, category-selective, response properties of the N170 is its larger amplitude and delayed latency following picture-plane inversion of a face stimulus (e.g., Rossion et al., 2000; Itier and Taylor, 2002). Over the last four decades, the N170 component has been described in hundreds, perhaps more than a thousand, of studies (Eimer, 2011; Rossion and Jacques, 2011 for reviews). In recent years, the N170 has been used as an index of deficient face recognition abilities in clinical populations, with moderate limited success (e.g., in Autism Spectrum Disorders (ASD): Batty et al., 2011; Jeffreys, 1989; see Jeffreys, 1996 for review). The N170 and VPP typically show similar response properties (Joyce and Rossion, 2005). One of the most salient, category-selective, response properties of the N170 is its larger amplitude and delayed latency following picture-plane inversion of a face stimulus (e.g., Rossion et al., 2000; Itier and Taylor, 2002). Over the last four decades, the N170 component has been described in hundreds, perhaps more than a thousand, of studies (Eimer, 2011; Rossion and Jacques, 2011 for reviews). In recent years, the N170 has been used as an index of deficient face recognition abilities in clinical populations, with moderate limited success (e.g., in Autism Spectrum Disorders (ASD): Batty et al., 2011; Jeffreys, 1989; see Jeffreys, 1996 for review). The N170 and VPP typically show similar response properties (Joyce and Rossion, 2005). One of the most salient, category-selective, response properties of the N170 is its larger amplitude and delayed latency following picture-plane inversion of a face stimulus (e.g., Rossion et al., 2000; Itier and Taylor, 2002).
Fig. 1. Grand average ERP waveforms elicited by different categories of stimuli at occipito-temporal electrode sites. Figures adapted from the studies of (A) Bentin et al. (1996), (B) Rossion and Jacques (2008), and (C) Rossion and Caharel (2011). The N170 is systematically larger in response to pictures of faces than to other categories and is practically absent for meaningless stimuli. This face-selectivity occurs well before the peak of the N170, at around 120–130 ms following stimulus onset, and is often limited to this time window (i.e., until 200 ms). (D) When pictures of the paintings of G. Arcimboldo are presented upright and usually lead to the perception of a face stimulus, a larger N170 is observed than when the exact same pictures are presented inverted and do not lead to the perception of a face (Figure adapted from Caharel et al., 2013). (E) The face-selectivity of the N170 is often clearly observed on ERP data of individual subjects (Figure adapted from Rousselet et al., 2008).
outstanding issue. While some ERP studies have reported long-term face familiarity effects (i.e., often amplitude increases) on the N170 (e.g., Caharel et al., 2005; Barragan-Jason et al., 2015), others have not reported such effects (e.g., Tanaka et al., 2006; Gosling and Eimer, 2011). Since these familiarity effects have been more consistently observed on subsequent ERP components (e.g., the N250; Schweinberger, et al., 2002a, 2002b; Tanaka et al., 2006; Wiese et al., 2019a, 2019b), the standard view in the scientific literature is that the N170 is an electrophysiological marker of a “perceptual stage” in face recognition (e.g., a “structural encoding stage” according to Bruce and Young, 1986’s influential model), which is followed by ERP components marking the access to representations of facial identities stored in memory (see Bentin et al., 1996; Bentin and Deouell, 2000; Eimer, 2000a, 2000b, 2011; Schweinberger and Neumann, 2016). From the point of view of fundamental research, clarifying this issue is therefore important because it can provide constraints on our understanding of the neurofunctional organization of human face recognition, i.e. how early in time is this function shaped by specific memories from our lifetime experience. Moreover, at least with a standard (i.e., slow and jittered) mode of visual stimulation to elicit ERPs, the N170 indexes the only reliable category-selective time-window (Fig. 1; Rossion and Jacques, 2008; Rousselet et al., 2008; Ganis et al., 2012), so that long-term familiarity effects observed during this window may be related to the contribution of face-selective populations of neurons. From the point of view of applied research, since the N170 is a very clear ERP, it could potentially serve as an implicit index to determine whether one knows a specific face identity or not.

SELECTION OF ERP STUDIES COMPARING FAMILIAR AND UNFAMILIAR FACES

In this brief review, we provide a list of all the published studies – to the best of our knowledge – that have directly contrasted the N170 elicited by pictures of unfamiliar and familiar faces in human adults. We also consider studies using magnetoencephalography (MEG), focusing on the M170 component (e.g., Kloth et al., 2006). Importantly, we do not consider studies that have compared the immediate repetition of a given face identity compared with the immediate presentation of different faces (i.e., face A followed immediately by face A or by face B) without directly comparing familiar to unfamiliar faces. For instance, the earliest ERP study testing immediate repetition of face identity reported an increase for different over repeated faces for familiar faces on a posterior ERP component corresponding to the N170 (Barrett et al., 1988). While this effect was found for familiar but not unfamiliar faces tested in the same repetition condition, indicating a genuine effect of familiarity at that latency, the authors did not directly compare familiar to unfamiliar faces. Moreover, we do not extend our review to face familiarity effects reported at different (i.e., typically later) time-windows in the EEG than the N170 (i.e., genuine post 200 ms effects, e.g., Wiese et al., 2019a, 2019b; or very early effects likely due to low-level physical differences between images, e.g., Seeck et al., 1997; Dobs et al., 2019). Finally, for the sake of fair comparison, our review focuses on studies reporting ERPs to adult faces only (e.g., excluding studies with children faces as seen by their own mothers; Grasso et al., 2009). However, our brief survey includes all studies contrasting responses to unfamiliar with familiar faces, such as those of famous and personally familiar persons, the participants’ own faces, as well as experimentally familiarized faces, and this regardless of the task used. It also considers studies using not only photographic faces but also pictures of faces manipulated experimentally (Mooney, caricatures, …). Only data from healthy adult participants with no sign of mental or neurological disorder are taken into account. Finally, only papers written in an English language are included.

Studies were selected from searches on the Pubmed and Google Scholar databases. The terms used for this search included (“N170” or “M170” “ERP” or “MEG”) and (“familiarity” or “familiar”). From the relevant articles from this research, a review of cited references was also carried out in order to find other articles. Overall, we identified 45 EEG or MEG studies that directly contrasted the N170 in response to familiar vs. unfamiliar faces. The following information were extracted from each of these studies and reported in tables (from 1 to 3) below: the number and mean age of participants, the task performed, the type of familiar reference electrode used for analyses.

THE KEY CONTRIBUTION OF PERSONALLY FAMILIAR FACES ON THE N170 FAMILIARITY EFFECT

Among all the studies included in the survey, twenty-one (46.7%) reported a significantly larger N170 amplitude to pictures of familiar faces (Table 1; Fig. 2) and twenty-one (46.7%) failed to report any significant difference (Table 2). Remarkably, only very few studies (3; 6.7%) found a significant difference in the other direction, i.e. a smaller N170 for familiar faces (Table 3).

What conclusions can be drawn from this brief survey?

First, although a familiarity effect on the N170 appears too small and inconsistent across studies to be of any use as an electrophysiological marker of face familiarity in applied research, too many studies, from various research groups, report a significant effect to still keep claiming that sensitivity to long-term familiarity does not emerge during the time-window of the N170 but only at later time points (see Fig. 2). A key element supporting this conclusion is the direction of the effect: if the modulation of the N170 by face familiarity was only due to EEG noise fluctuations, there would be no reason to expect a consistent direction of this amplitude difference across studies, i.e. in the form of larger amplitude for familiar faces (i.e., familiar > unfamiliar: 21/24 studies; 87.5%; considering only the studies showing a
significant difference: 46.7% vs. 6.7%, $X^2 = 18.19$, $p < .001$). Importantly, since ERP studies typically focus on multiple components, there is also no reason to think that only those with positive face familiarity N170 effects are published, artificially increasing the proportion of positive results identified in the present review. If

Table 1. List of studies that reported a larger N170 amplitude for familiar than unfamiliar faces

<table>
<thead>
<tr>
<th>Authors</th>
<th>Year</th>
<th>N (controls) (mean age)</th>
<th>Task</th>
<th>Familiar face types</th>
<th>Number of faces per familiarity level</th>
<th>Number of repetitions for each face</th>
<th>Reference electrode used for analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caharel et al.</td>
<td>2002</td>
<td>11 (24.5)</td>
<td>Passive view</td>
<td>Own face/Famous</td>
<td>1</td>
<td>100</td>
<td>Common average</td>
</tr>
<tr>
<td>Herzmann et al.</td>
<td>2004</td>
<td>16 (20.9)</td>
<td>Familiarity judgment (priming)</td>
<td>Personally familiar (lecturing staff of the University)/Famous</td>
<td>15 Personally familiar/15 Famous/30 Unfamiliar</td>
<td>6</td>
<td>Common average</td>
</tr>
<tr>
<td>Caharel et al.</td>
<td>2005</td>
<td>14 (23)</td>
<td>Task 1: Familiarity judgment/Task 2: Expression judgment</td>
<td>Own face/Personally familiar (mother)/Famous</td>
<td>1 Own and Personally familiar/20 Famous and Unfamiliar</td>
<td>100/5</td>
<td>Common average</td>
</tr>
<tr>
<td>Mnatsakanian and Tarkka</td>
<td>2005</td>
<td>13 (nr)</td>
<td>Face and Pattern (dots) tasks</td>
<td>Personally familiar</td>
<td>nr</td>
<td>nr</td>
<td>Common average</td>
</tr>
<tr>
<td>Caharel et al.</td>
<td>2006</td>
<td>20 (24)</td>
<td>Familiarity judgment</td>
<td>Personally familiar (family members and friends)</td>
<td>10</td>
<td>10</td>
<td>Common average</td>
</tr>
<tr>
<td>Kloth et al. -- MEG</td>
<td>2006</td>
<td>12 (24.8)</td>
<td>Familiarity judgment</td>
<td>Personally familiar (students and lecturers)/Famous</td>
<td>66 Personally familiar/66 Famous/132 Unfamiliar</td>
<td>2</td>
<td>--</td>
</tr>
<tr>
<td>Caharel et al.</td>
<td>2007</td>
<td>18 (36.3)</td>
<td>Familiarity judgment</td>
<td>Own face, Personally familiar (University professors)</td>
<td>1 Own and Personally familiar/25 Unfamiliar</td>
<td>50/25</td>
<td>Common average</td>
</tr>
<tr>
<td>Zeman et al.</td>
<td>2007</td>
<td>12 (31)</td>
<td>Target (Joe face) decision</td>
<td>Own face/Experimentally learned (Joe or Jane)</td>
<td>1 Own/1 Learned/100 Unfamiliar/170 Own/70 Learned/7 Unfamiliar</td>
<td>2</td>
<td>--</td>
</tr>
<tr>
<td>Dobel et al. -- MEG</td>
<td>2008</td>
<td>7 (39)</td>
<td>Familiarity judgment</td>
<td>Famous</td>
<td>66</td>
<td>2</td>
<td>--</td>
</tr>
<tr>
<td>Wild-Wall et al.</td>
<td>2008</td>
<td>Task 1: 16 (24.6)/Task 2: 20 (25.0)</td>
<td>Task 1: Expression judgment/Task 2: Familiarity judgment</td>
<td>Personally familiar (staff members of the University)</td>
<td>14</td>
<td>Task 1: 12/Task 2: 6</td>
<td>Common average</td>
</tr>
<tr>
<td>Harris and Aguirre -- MEG</td>
<td>2008</td>
<td>14 (18–35)</td>
<td>Target (flower) detection</td>
<td>Famous</td>
<td>40</td>
<td>2</td>
<td>--</td>
</tr>
<tr>
<td>Keyes et al.</td>
<td>2010</td>
<td>19 (23.4)</td>
<td>Repetition detection</td>
<td>Own face/Personally familiar (friends)</td>
<td>1</td>
<td>80</td>
<td>CMS/DRL</td>
</tr>
<tr>
<td>Leleu et al.</td>
<td>2010</td>
<td>15 (23)</td>
<td>Switching between Familiarity and Expression judgments</td>
<td>Famous</td>
<td>20</td>
<td>16</td>
<td>Common average</td>
</tr>
<tr>
<td>Caharel et al.</td>
<td>2011</td>
<td>20 (17.9)</td>
<td>Same-different identity matching (adaptation)</td>
<td>Personally familiar (students)</td>
<td>23</td>
<td>3</td>
<td>Common average</td>
</tr>
<tr>
<td>Geng et al.</td>
<td>2012</td>
<td>15 (23.1)</td>
<td>Fixation change detection in supraliminal or subliminally condition</td>
<td>Own face/Famous</td>
<td>1</td>
<td>nr</td>
<td>Common average</td>
</tr>
<tr>
<td>Caharel et al.</td>
<td>2014</td>
<td>22 (23)</td>
<td>Go/no-go familiarity judgment</td>
<td>Personally familiar (students)</td>
<td>26</td>
<td>8</td>
<td>Common average</td>
</tr>
<tr>
<td>Alonso-Prieto et al.</td>
<td>2015</td>
<td>10 (28.1)</td>
<td>Pleasant or unpleasant face judgment</td>
<td>Famous</td>
<td>5</td>
<td>40</td>
<td>Common average</td>
</tr>
<tr>
<td>Barragan-Jason et al.</td>
<td>2015</td>
<td>28 (24)</td>
<td>Go/no-go familiarity judgment</td>
<td>Famous</td>
<td>270</td>
<td>0</td>
<td>Common average</td>
</tr>
<tr>
<td>MacKenzie and Donaldson</td>
<td>2016</td>
<td>Exp 1: 20 (21)/Exp 2: 24</td>
<td>Familiar judgment with greyscale/cropped faces (exp1) or color/native background (exp2)</td>
<td>Famous</td>
<td>Exp 1: ~50/Exp 2: 100</td>
<td>1</td>
<td>Common average</td>
</tr>
<tr>
<td>Podvigina and Prokopenya</td>
<td>2019</td>
<td>26 (19–35)</td>
<td>Familiarity judgment of words or faces</td>
<td>Famous</td>
<td>80</td>
<td>1</td>
<td>Linked-earlobes</td>
</tr>
<tr>
<td>Wiese et al.</td>
<td>2019b</td>
<td>Exp 1: 18 (21.7)/Exp 2: 18</td>
<td>Letter string and butterfly detection with high (exp 1) or low (exp 2) resource demands</td>
<td>Personally familiar (close friends)</td>
<td>1</td>
<td>100</td>
<td>Common average</td>
</tr>
</tbody>
</table>
anything, as discussed below, the opposite bias might be present because researchers typically expect effects of face familiarity to emerge on later ERP components, so that effects of face familiarity on the N170 are not always well reported and evaluated (e.g., Herzmann et al., 2004; Wiese et al., 2019a).

Second, given the wide variety of paradigms and the lack of methodological information reported in some studies, it is, unfortunately, almost impossible to identify clear factors (e.g., expectations/task, number of images, image repetition, time-windows selected for analysis, etc.) explaining the presence or absence of a familiarity

Fig. 2. Grand average ERP waveforms elicited by familiar (famous, personally familiar, and own faces) and unfamiliar faces at right occipito-temporal electrodes. Figures adapted from the studies of (A) Caharel et al. (2006), (B) Keyes et al., (2010), (C) Caharel et al., (2011), (D) Barragan-Jason et al. (2015) and (E) Wiese et al. (2019b). The N170 amplitude is larger for familiar than for unfamiliar faces and this difference occurs at its peak (i.e., at around 170 ms), that is to say after the face-selectivity effect emerging at the onset of the N170 (i.e., 120–130 ms).
Table 2. List of studies that failed to report a N170 difference in amplitude between familiar and unfamiliar faces

<table>
<thead>
<tr>
<th>Authors</th>
<th>Year</th>
<th>N (controls) (mean age)</th>
<th>Task</th>
<th>Number of faces per familiarity level</th>
<th>Number of repetitions for each face</th>
<th>Reference electrode used for analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rossion et al.</td>
<td>1999</td>
<td>9 (19–26)</td>
<td>Task 1: Gender discrimination/Task 2: Familiarity judgment</td>
<td>Experimentally learned</td>
<td>20 Learned/40 Unfamiliar</td>
<td>Common average</td>
</tr>
<tr>
<td>Bentin and Deouell</td>
<td>2000</td>
<td>Task 1: 14 (nr)/Task 2: 24 (nr)</td>
<td>Task 1: to count butterflies/Task 2: to count famous politicians</td>
<td>Famous</td>
<td>60 0</td>
<td>Tip of the nose</td>
</tr>
<tr>
<td>Eimer</td>
<td>2000a</td>
<td>15 (24.9)</td>
<td>Task 1: target (hands) detection/Task 2: target (digits) detection or repetition detection</td>
<td>Famous</td>
<td>50 4</td>
<td>Tip of the nose</td>
</tr>
<tr>
<td>Eimer</td>
<td>2000b</td>
<td>24 (27)</td>
<td>Task 1: target (hands) detection/Task 2: target (digits) detection or repetition detection</td>
<td>Famous</td>
<td>50 6</td>
<td>Tip of the nose</td>
</tr>
<tr>
<td>Schweinberger et al.</td>
<td>2002a</td>
<td>12 (21.2)</td>
<td>Familiarity judgment (priming)</td>
<td>Famous</td>
<td>90 0</td>
<td>Common average</td>
</tr>
<tr>
<td>Schweinberger et al.</td>
<td>2002b</td>
<td>18 (22.7)</td>
<td>Familiarity judgment (priming)</td>
<td>Famous</td>
<td>90 0</td>
<td>Common average</td>
</tr>
<tr>
<td>Henson et al.</td>
<td>2003</td>
<td>18 (23)</td>
<td>Task 1: left–right symmetry judgment/Task 2: gender categorization</td>
<td>Famous</td>
<td>64 0</td>
<td>Mastoids</td>
</tr>
<tr>
<td>Engst et al.*</td>
<td>2006</td>
<td>16 (25.6)</td>
<td>Familiarity judgment (priming)</td>
<td>Famous</td>
<td>64 1</td>
<td>Common average</td>
</tr>
<tr>
<td>Sui et al.</td>
<td>2006</td>
<td>18 (24.7)</td>
<td>Head orientation judgment</td>
<td>Own face/Personally familiar (classmates or roommates)</td>
<td>1 160</td>
<td>Mastoids</td>
</tr>
<tr>
<td>Tanaka et al.</td>
<td>2006</td>
<td>24 (20.79)</td>
<td>Target (Joe face) decision</td>
<td>Own face/Experimentally learned (Joe or Jane)</td>
<td>1 70</td>
<td>Common average</td>
</tr>
<tr>
<td>Anaki et al.</td>
<td>2007</td>
<td>24 (25)</td>
<td>Task 1: Target (flower) detection/Task 2: Familiarity judgment</td>
<td>Famous</td>
<td>90 0</td>
<td>Tip of the nose</td>
</tr>
<tr>
<td>Ewbank et al. MEG</td>
<td>2008</td>
<td>18 (23)</td>
<td>Target (red dot) detection task (adaptation)</td>
<td>Famous</td>
<td>13 208</td>
<td>_</td>
</tr>
<tr>
<td>Miyakoshi et al.</td>
<td>2008</td>
<td>16 (20.4)</td>
<td>Familiarity judgment</td>
<td>Own face/Famous</td>
<td>1 50</td>
<td>Common average</td>
</tr>
<tr>
<td>Kaufmann et al.</td>
<td>2009</td>
<td>24 (22.9)</td>
<td>Familiarity judgment (priming)</td>
<td>Experimentally learned</td>
<td>70 Learned/140 Unfamiliar</td>
<td>Common average</td>
</tr>
<tr>
<td>Dörr et al.*</td>
<td>2011</td>
<td>16 (26.4)</td>
<td>Familiality judgment (priming)</td>
<td>Famous</td>
<td>116 1</td>
<td>Common average Linked-earlobes</td>
</tr>
<tr>
<td>Gosling and Eimer</td>
<td>2011</td>
<td>16 (27.3)</td>
<td>Familiality judgment</td>
<td>Famous</td>
<td>80 4</td>
<td>Common average</td>
</tr>
<tr>
<td>Pierce et al.</td>
<td>2011</td>
<td>12 (27.25)/Task 2: 24 (nr)</td>
<td>Task 1: Target (Joe or his dog) decision/Exp 2: Target (Joe or his car) decision</td>
<td>Own face/Experimentally learned (Joe or Jane)</td>
<td>1 Own/1 Learned/4 Unfamiliar</td>
<td>Task 1: Linked-earlobes/Task 2: Common average</td>
</tr>
<tr>
<td>Schulz et al.</td>
<td>2012</td>
<td>31 (22.9)</td>
<td>Two-choice (learned or not learned) task (3 distinctiveness conditions)</td>
<td>Experimentally learned</td>
<td>30 3</td>
<td>Common average</td>
</tr>
<tr>
<td>Andrews et al.</td>
<td>2017</td>
<td>19 (21.95)</td>
<td>Target (butterfly) detection</td>
<td>Famous/Experimentally learned</td>
<td>2 Famous/ 6 Learnd/ nr</td>
<td>Common average</td>
</tr>
<tr>
<td>Alzueta et al.</td>
<td>2019</td>
<td>25 (22.7)</td>
<td>Familiarity judgment</td>
<td>Own face/Personaly familiar (friends)</td>
<td>1 150</td>
<td>Common average</td>
</tr>
<tr>
<td>Wiese et al.*</td>
<td>2019a</td>
<td>20 (20.7)/Exp 3: 18 (19.9)</td>
<td>Target (butterfly) detection</td>
<td>Exp 1: Highly personally familiar (relatives) and in Exp 3: less personally familiar (lecturer) and Famous</td>
<td>1 50</td>
<td>Common average</td>
</tr>
</tbody>
</table>

("nr" for "not reported"; * studies showing trends towards significance).
effect on the N170 (Table 1). This is also a major reason why a systematic meta-analysis on this issue is currently impossible. For instance, the nature of the task (explicit vs. implicit familiarity judgment) does not seem to explain the divergence of results across the studies. Indeed, an explicit task was used in 13 of the 21 studies that reported a N170 face familiarity effect (62%, Table 1) and in 12 of 21 studies (57%, Table 2) which failed to report this effect ($\chi^2 = 0.106$, $p = .74$). Also, about the same proportion of studies that showed (15/21; 71%) or did not show (12/21 = 57%) an effect, used the recommended (Joyce and Rossion, 2005) common average reference ($\chi^2 = 0.87$, $p = .35$), suggesting that this factor is not directly relevant. Moreover, the same number of studies that showed (7/21; 33%) or did not show (7/21, 33%) a familiarity effect used a number of repetitions of each face identity greater than 40. Yet, interestingly, the level of familiarity, in particular the personal familiarity with the faces used as stimuli, seems to play an important role. First, the majority (14/21; 67%) of studies using pictures of personally familiar faces report a larger N170 to familiar than unfamiliar faces (e.g., Caharel et al., 2002, 2005, 2006, 2007, 2011, 2014; Herzmann et al., 2004; Mnatsakanian and Tarkka, 2005; Kloth et al., 2006; Zeman et al., 2007; Wild-Wall et al., 2008; Keyes et al., 2010; Geng et al., 2012; Wiese et al., 2019b; Table 1), with only six studies (29%) that did not report such effect (67% vs. 29%, $\chi^2 = 5.93$, $p = .015$). Moreover, all of the six ERP studies that did not report any N170 modulation for personally familiar faces used a single face identity (often participants’ own face) repeated dozens of times (Sui et al., 2006; Tanaka et al., 2006; Miyakoshi et al., 2008; Pierce et al., 2011; Alzueta et al., 2019; Wiese et al., 2019a) (Table 2). Most recently, Wiese et al. (2019a) contrasted pictures of two personally (highly) familiar face identities to two unfamiliar face identities, but with variable pictures for both of them. In the main (#1) experiment of that study, the N170 was larger in amplitude for personally familiar faces but the difference did not reach significance ($p = .07$; supplemental material of Wiese et al., 2019a). However, a careful look at the results of this study strongly suggests in fact an effect of familiarity on the N170 (bottom waveforms on Fig. 2a of Wiese et al., 2019a). Interestingly, while the interaction with the factor block – for stimulus repetition – was not reported by the authors, the waveforms of the last two blocks at least clearly show a larger N170 amplitude for familiar faces (see also Fig. 2a of Wiese et al., 2019a). This is consistent with a subsequent study of these authors reporting significant face familiarity effect on the N170 (Wiese et al., 2019b; see Fig. 2), and suggesting overall that the N170 is genuinely larger in amplitude for pictures of personally familiar faces than unfamiliar faces, even if this effect is not reported or emphasized in the published studies.

Why is the N170 familiarity effect found more often for pictures of personally familiar faces than famous or familiarized faces? On the one hand, personally familiar faces are encountered much more regularly and are learned in more natural and varied visual conditions than famous faces. They may be automatically associated with more, and more diverse, semantic and affective traits, making their neural representations more “robust”, facilitating detection, recognition of identity and social cues, and activation of person knowledge (Ramon and Gobbini, 2018). On the other hand, it is also likely that participants of a given study diverge greatly in their knowledge of celebrities, and that their level of familiarity for various famous faces also varies substantially, potentially blurring any N170 effect of long-term familiarity for famous faces. Another factor could be that participants are familiar with the specific pictures of the personally familiar faces used in an experiment but not with those of famous faces, potentially boosting the effects observed for the former. However, this may be an issue for famous faces too, which are often presented under well-known ‘iconic’ views (Carbon, 2008).

Third, the N170 is a relatively early response originating from high-level visual areas of the human brain, and is therefore certainly more sensitive than later components (e.g., N250) to physical aspects of stimuli. Hence, any imbalance between the physical aspects of image sets compared in a given study is likely to affect more a difference at the level of the N170 than at following components, potentially blurring an early effect of face familiarity. Unfortunately, physical aspects of familiar and unfamiliar face sets are rarely if ever described and controlled in the published studies. However, across studies, consistent effects of long-term personal face familiarity on the N170 are unlikely to reflect image-based effects because, contrary to sets of famous faces used, personally familiar faces differ for every participant of a given study.

Fourth, while face-selectivity effects (i.e., differences between faces and other visual stimuli) emerge generally at the onset of the N170 time window (i.e., 120–130 ms, see Rossion and Jacques, 2008), positive face familiarity effects are rather observed on the peak
of the component (at around 170 ms) (Fig. 2), i.e. indicating a delay of at least 40–50 ms between the two types of responses. This indicates that recognizing a stimulus as a face temporally precedes the recognition of its familiarity as an individual, but that the two functions nevertheless largely overlap in time, within the N170 time window. To our knowledge, a relationship between these two functions, in terms of EEG signal, across individuals, has not been established and could be further investigated in future studies. Moreover, interestingly, effects of familiarity appear to emerge at the same time as the (immediate) face identity repetition effect for unfamiliar faces (i.e., also at the peak of the N170), this latter effect being observed across substantial changes of size and head orientation but not stimulus inversion (i.e., ruling out an effect merely due to low-level visual cues; Jacques et al., 2007; Caharel et al., 2009; Caharel et al., 2015).

Fifth, a familiarity effect on the N170 does not mean that all that the process of attributing familiarity to a face stimulus is completed at this level, but only that enough evidence has been accumulated by the visual system to elicit an early detectable (i.e., significant) response on scalp EEG. In contrast, a much larger and consistent prolonged face familiarity recognition response from 200 ms onset has been described in numerous studies (i.e., on a N250 component; e.g., Schweinberger et al., 2002a, 2002b; Herzmann et al., 2004; Tanaka et al., 2006; Pierce et al., 2011; Wiese et al., 2019a, 2019b) and may better serve this purpose. In the same vein, a face familiarity effect does not mean that access to the unique familiar identity of the face has been completed at this level (after all, how often do we recognize that a face is familiar without being able to identify it, let alone recall its name?). However, recognizing (i.e., providing a selective reproducible response to) a face as being familiar implies that it is at least differentiated at the individual level from unfamiliar faces. That is, the distinction between a familiar and an unfamiliar face requires access to representations of specific individuals derived from past experience.

THEORETICAL IMPLICATIONS AND FUTURE DIRECTIONS

Following this brief synthesis, a number of outstanding questions or implications can also be raised for future studies. For instance, since it is well established that increased N170 amplitude depends on the subjective experience of a face, i.e. perceptual awareness (Srebro, 1985; Caharel et al., 2013; Navajas et al., 2013; Churches et al., 2014; see Rossion, 2014 for review), does this also apply to the familiarity effect observed as early as this component? In other words, is the familiarity effect (typically an amplitude increase) of the N170 associated with awareness of perceiving a familiar face or could it be elicited in the absence of conscious recollection (i.e., a form of covert familiar face recognition; Bruyer et al., 1983; De Haan et al., 1991)?

Most importantly, a genuine effect of familiarity on the N170, i.e. starting at an earlier level than expected, appears to challenge standard (hierarchical) functional models of human face recognition (Bruce and Young, 1986; Calder and Young, 2005; Young and Bruce, 2011). Indeed, according to this standard model, the N170 has been associated with a “structural encoding stage” (e.g., Bentin et al., 1996; Bentin and Deouel, 2000; Eimer, 2000a, 2000b, 2011). That is, a visuo-perceptual processing stage at which the idiosyncratic characteristics of the face are exacted, thought to occur before and independently of the activation of representations of faces stored in memory. Later ERP components, such as the N250 or the N400, would then reflect the association of this perceptual representation to representations of faces stored in memory and access to specific semantic information (i.e., familiar faces) (Bentin and Deouel, 2000; Eimer, 2000a, 2000b; Schweinberger et al., 2002a, 2002b; Tanaka et al., 2006; Pierce et al., 2011; Saavedra et al., 2012; Towler et al., 2012; Olivares et al., 2018; Schulz et al., 2012).

In light of the present review, this standard functional model may still hold, but the structural encoding stage for faces may already be affected by previous experience with specific facial identities and thus not be purely based on physical features of the stimulus (e.g., in a “face-space” representation; Valentine et al., 2016). This would be compatible with the well-documented observations that familiar faces are matched much better according to their identity across different viewing conditions than unfamiliar faces (e.g., Bruce, 1982; Jenkins et al., 2011), although this effect could also well be due to the rich associations of familiar faces with semantic, affective and even verbal information (Bruce, 1982; Schwartz and Yovel, 2016; Rossion, 2018). In any case, an adjustment/revision of the two-stage (perceptual/mnemonic) standard functional model of human face recognition and this association with successive ERP components would be required.

Alternatively, this distinction between two independent processing stages, as well as their respective association with consecutive ERP components, may have to be more fundamentally revised. According to an alternative theoretical framework, there would not be “perceptual stage” preceding recognition, perception being rather defined as the subjective experience that arises from recognition (Rossion and Retter, 2020). Recognition of faces occurs when (low-level, topographically organized, unconscious) sensory inputs match representations of faces in memory (i.e., populations of neurons in the human ventral-occipito-temporal cortex that have been selected to respond to specific inputs across different viewing conditions). According to this latter view, rather than reflecting two subsequent stages with two levels of facial representations, the successive effects observed in EEG would merely reflect accumulation of sensory evidence within the same neuro-functional system. That is, within the same selective populations of neurons in the ventral-occipito-temporal cortex, there is sufficient accumulation of evidence to significantly distinguish between faces and nonface categories (i.e., generic face recognition) precociously (i.e., at the N170 onset, at about 120–130 ms, Rossion and Jacques, 2008), with additional evidence required for the finer-grained distinction between...
familiar and unfamiliar face identities (i.e., face identity recognition and familiarity effects beginning at the N170 peak, at around 160–170 ms). The familiarity effect starting at about the N170 peak latency would simply reflect the onset of a prolonged single continuous process of (long-term) face familiarity, which would be expressed prominently during the N250 time-window, i.e. until about 400 ms. The richer the representation of familiar faces, as for instance in the case of personally familiar faces, the earlier and larger the differential neural response signaling familiarity.

This alternative theoretical account is certainly more parsimonious because it does not require two copies/representations of face identities (i.e., a perceptual one, derived from the input, and a representation stored in memory) (Rossion and Retter, 2020). It also agrees with the lack of clear-cut division between apperceptive and associative forms of prosopagnosia (Farah, 1990/2004), and the lack of reliable differences between neural sources of the N170 and N250, both thought to originate mainly from the (right) ventral occipito-temporal cortex (Schweinberger et al., 2002b; Kaufmann et al., 2009; Rossion and Jacques, 2011; Olivares et al., 2018). Finally, it also incorporates the well-documented effects of short-lagged immediate unfamiliar face repetition also starting on the N170 peak (Rossion and Jacques, 2011 for review) and extending until about the 200–400 ms time-window, into an integrated theoretical framework.

Evaluating this proposal with human electrophysiological recordings certainly requires going beyond description of ERP components as stages of processing and focus on differential waveforms between conditions (with confidence intervals), avoiding the problem of a subjective and often suboptimal selection of time-windows of interest (e.g., 140–180 ms) for data analysis. In recent years, human EEG studies with face stimuli have been increasingly performed using a fast periodic mode of visual stimulation, generating high signal-to-noise ratio (SNR) responses in the frequency-domain exactly at the stimulation frequency rate (‘steady-state visual evoked potentials’ or ‘frequency-tagging’, see Norcia et al., 2015 for review). This type of approach holds significant advantages in terms of sensitivity and objectivity, and the signal can be compared for familiar and unfamiliar faces presented during different fast periodic trains of images (Collins et al., 2018). Most interestingly, pictures of familiar faces can be periodically (e.g., 1.2 Hz) inserted in rapid (6 Hz) periodic trains of unfamiliar faces (Zimmermann et al., 2019). The common response to familiar and unfamiliar faces can therefore be selectively projected to the general stimulation frequency (6 Hz), allowing to focus on their differential response expressed at the slower familiar face stimulation frequency (Zimmermann et al., 2019; Campbell et al., 2020; Yan et al., 2020). These studies have generally concentrated on frequency-domain EEG responses, reporting robust differences between familiar and unfamiliar face identities over the occipito-temporal cortex in individual participants (Zimmermann et al., 2019; Campbell et al., 2020; Yan et al., 2020). Yet, since time-domain analyses of such effects can also be performed, providing that adequate stimulation parameters are used (see Retter and Rossion, 2016; Rossion et al., 2020), they should be able to reveal robust differential waveforms between familiar and unfamiliar faces emerging at around 170 ms as sustained effects rather than separate components reflecting different functional ‘stages’ (see Yan and Rossion, 2020).

Finally, while recordings of face-selective responses inside the human brain have been described for decades with intracranial recordings (Allison et al., 1994; Halgren et al., 1994), there has been a surge of such studies in recent years, with recordings performed inside brain regions typically defined as face-selective in fMRI such as the lateral section of the (right) middle fusiform gyrus (Jacques et al., 2016; Jonas et al., 2016). The present overview suggests that effects of long-term familiarity of the faces – especially for personally familiar faces – should be observed at the level of intracranial responses, most likely within face-selective occipito-temporal regions, emerging significantly before 200 ms.

In conclusion, our brief survey of two decades of ERP studies in which pictures of long-term familiar and unfamiliar faces were compared indicates that a significant effect linked to familiarity with specific facial identities emerges earlier in the human brain than generally thought, i.e. between 150–200 ms in occipito-temporal brain regions. This effect briefly follows the emergence of reliable face-selectivity at the N170 onset, but occurs at the same time as the earliest high-level effects of immediate unfamiliar face identity repetition. These observations challenge standard neurocognitive models with a clear-cut distinction between perceptual and memory stages in human face recognition.

ACKNOWLEDGMENTS

The authors would like to thank John Towler, Holger Wiese and an anonymous reviewer for their helpful comments on a previous version of this manuscript.

REFERENCES


Received 3 August 2020. Accepted 31 December 2020
(Available online xxxx)