ACCEPTED AUTHOR VERSION, TO BE PUBLISHED IN *PSYCHOPHYSIOLOGY*

Detecting a viewer's familiarity with a face:

Evidence from event-related brain potentials and classifier analyses

Holger Wiese[1], Dasha Anderson[1], Ulrik Beierholm[1], Simone C. Tüttenberg[2],

Andrew W. Young[3], & A. Mike Burton[3]

[1]*Department of Psychology, Durham University, UK,*

[2]*Department of Psychology, Saarland University, Germany, and*

[3]*Department of Psychology, University of York, UK*

Corresponding author:

Holger Wiese, Department of Psychology, Durham University, South Road, Durham DH1 3LE, United Kingdom. E-Mail: holger.wiese@durham.ac.uk

Abstract

Human observers recognise the faces of people they know efficiently and without apparent effort. Consequently, recognising a familiar face is often assumed to be an automatic process beyond voluntary control. However, there are circumstances in which a person might seek to hide their recognition of a particular face. The present study therefore used event-related potentials (ERPs) and a classifier based on logistic regression to determine if it is possible to detect whether a viewer is familiar with a particular face, regardless of whether the participant is willing to acknowledge it or not. In three experiments, participants were presented with highly variable "ambient" images of personally familiar and unfamiliar faces, while performing an incidental butterfly detection task (Experiment 1), an explicit familiarity judgment task (Experiment 2), and a concealed familiarity task in which they were asked to deny familiarity with one truly known facial identity while acknowledging familiarity with a second known identity (Experiment 3). In all three experiments, we observed substantially more negative ERP amplitudes at occipito-temporal electrodes for familiar relative to unfamiliar faces starting approximately 200 ms after stimulus onset. Both the earlier N250 familiarity effect, reflecting visual recognition of a known face, and the later Sustained Familiarity Effect (SFE), reflecting the integration of visual with additional identity-specific information, were similar across experiments and thus independent of task demands. These results were further supported by the classifier analysis. We conclude that ERP correlates of familiar face recognition are largely independent of voluntary control and discuss potential applications in forensic settings.

Keywords: Face Recognition, Voluntary Control, Event-Related Potentials, Sustained Familiarity Effect, MVPA

1. Introduction

Human observers are highly efficient at recognising the faces of the people they know (e.g., Bruce & Young, 2012) and accomplish this task many times every day without apparent effort. Yet the complexity of the task becomes evident when comparing familiar to unfamiliar face recognition (Young & Burton, 2017). It is usually easy to see that two or more very different images show the same familiar face but recognising that different images show the same unfamiliar face is often surprisingly difficult (Bruce et al., 1999; Jenkins, White, Van Montfort, & Burton, 2011). Consequently, both naïve participants and trained professionals can have substantial difficulties identifying an unfamiliar person on an identity card photo (Kemp, Towell, & Pike, 1997; White, Kemp, Jenkins, Matheson, & Burton, 2014) or from CCTV footage (Bruce, Henderson, Newman, & Burton, 2001; Henderson, Bruce, & Burton, 2001).

Familiar face recognition is often held to be largely automatic (Bruce & Young, 2012; for very recent neural evidence, see Dalski, Kovacs, & Ambrus; Yan, Young, & Andrews, 2017; Young & Burton, 2018). While there are numerous definitions of the term, most agree that automatic processes (i) are fast, and do not depend on (ii) conscious awareness, (iii) attentional resources, or (iv) intention or voluntary control (Logan, 1988; Moors, 2016; Posner & Snyder, 1975; Shiffrin & Schneider, 1977). In particular, lack of voluntary control is evident from the commonplace observation that you cannot look at a familiar face and decide not to recognise it. Previous research has also demonstrated that familiar face recognition can indeed occur without conscious awareness, as implicit forms have been observed in the absence of explicit face recognition (Bauer, 1984; Tranel, Damasio, & Damasio, 1995; Young, Hellawell, & De Haan, 1988). Moreover, familiar face recognition does not seem to depend on the availability of general attentional resources (Jenkins, Burton,

& Ellis, 2002; Lavie, Ro, & Russell, 2003), at least when processing is limited to one as opposed to multiple faces at a time (Bindemann, Burton, & Jenkins, 2005).

Although familiar face recognition may itself be automatic and therefore involuntary, there are circumstances in which a perceiver might be motivated to seek to deny having recognised a particular face; for example, in a police investigation if they are both members of a criminal organisation. In the studies presented here, we therefore examine the extent to which automatic aspects of familiar face recognition can be determined from EEG measures with high temporal sensitivity. Complementary to this, we manipulate the task performed by viewers to estimate the extent to which they can 'hide' their ability to recognise a face from the EEG-based measures. In Experiment 1, participants passively view familiar and unfamiliar faces, while responding only to images from a different category (butterfly detection); this incidental task does not require explicit recognition of the familiar faces. In Experiment 2, we ask participants to make familiarity judgements to each of a sequence of face photos, thus demanding explicit recognition of familiar identities. In Experiment 3, we instruct participants to make familiarity judgements to faces, but sometimes to conceal their true knowledge by responding ‘unfamiliar’ to a face they actually know; this task requires additional processing to counteract a response based on automatic recognition. Across these experiments we record event-related brain potentials (ERPs) and examine components known to be sensitive to the recognition of familiar faces. The modulation of these components by the viewers’ task offers direct insights into the automaticity of familiar face recognition. In particular, a lack of modulation by task would provide evidence that ERPs could be used to judge the familiarity of a face even in cases where the viewer is trying to conceal it – thus providing a potentially useful tool in forensic examination.

The studies presented below extend previous research in a number of ways. First, we aim to examine the recognition of personally familiar faces (rather than celebrities) by tailoring stimuli to individual participants. Second, we examine familiarity using multiple photos of the same person, rather than repetition of single images. Hence, we are focussed here specifically on *identity* recognition, rather than *image* recognition, and this distinction is specifically examined in Experiment 1. Third, we design the data analysis to reveal individual-level as well as group effects, with both classic ERPs and an alternative approach based on a classifier using logistic regression. If the techniques described are to be used in forensic settings, it is important to understand their individual reliability, as well as the overall effects across experimental samples. We next outline the details of our approach.

1.1 ERPs and face recognition

Face and person recognition are typically conceptualised as a sequence of processing steps (Bruce & Young, 1986; Gobbini & Haxby, 2007; Schweinberger & Neumann, 2016). The present series of experiments therefore investigated ERP correlates of familiar face recognition, which allows the more isolated examination of these sub-processes (for a review, see Olivares, Iglesias, Saavedra, Trujillo-Barreto, & Valdes-Sosa, 2015). The first face-sensitive ERP component, the N170, peaks approximately 170 ms after stimulus onset at occipito-temporal channels and clearly distinguishes between faces and other visual objects (Bentin, Allison, Puce, Perez, & McCarthy, 1996; Eimer, 2011; Rossion & Jacques, 2008). While some studies have found familiarity effects in the N170 (Caharel, Courtay, Bernard, Lalonde, & Rebai, 2005; Caharel, Jacques, d'Arripe, Ramon, & Rossion, 2011; Johnston, Overell, Kaufman, Robinson, & Young, 2016), differences between familiar and unfamiliar faces are more consistently observed in later time windows.

From approximately 200 ms after stimulus onset, familiar faces elicit more negative amplitudes relative to unfamiliar faces at occipito-temporal electrodes. This so-called N250 familiarity effect has been observed in studies contrasting both celebrity (Bentin & Deouell, 2000; Gosling & Eimer, 2011; Saavedra, Iglesias, & Olivares, 2010) and newly learnt with unfamiliar faces (Andrews, Burton, Schweinberger, & Wiese, 2017; Kaufmann, Schweinberger, & Burton, 2009; Tanaka, Curran, Porterfield, & Collins, 2006). Importantly, the N250 effect has been observed when participants are substantially distracted from the face stimuli (see also Neumann & Schweinberger, 2008; Wiese, Ingram, et al., 2019), and even in the absence of explicit recognition in developmental prosopagnosia (Eimer, Gosling, & Duchaine, 2012). Given its occurrence in the first few hundred milliseconds of stimulus processing, these findings suggest that the N250 fulfils the first three criteria of automaticity (occurring fast and independent of conscious awareness and attentional resources).

Recently, we described an additional ERP familiarity effect (Wiese, Tüttenberg, et al., 2019). Participants were presented with multiple task-irrelevant ambient images of highly personally familiar and unfamiliar faces while responding only to randomly intermixed pictures of butterflies. We observed substantially more negative ERP amplitudes for familiar than unfamiliar faces at occipito-temporal electrodes between 400-600 ms, and, although highly similar in scalp distribution, this Sustained Familiarity Effect (SFE) was substantially larger than the earlier N250 effect. We tentatively interpreted the SFE as reflecting the integration of visual with additional identity-specific (e.g. semantic, episodic, affective) information. In contrast to the N250 effect, the SFE is reduced when participants are distracted from the face stimuli (Wiese, Ingram, et al., 2019), suggesting that the effect depends on the availability of attentional resources. It therefore does not seem to fulfil the corresponding criterion of automaticity, while the role of voluntary control for the SFE is unclear.

1.2 Detecting concealed knowledge

Numerous ERP studies have examined the possibility of detecting information the participant is trying to hide from the investigator using so-called Guilty Knowledge or Concealed Information Tests (CIT). In a typical experiment, details of staged “crimes” that are only known to a “guilty” person (such as a specific weapon used in the crime or objects present at the crime scene) are presented among crime-irrelevant, but otherwise comparable information. Participants are typically asked to respond to the presentation of target items, which are defined by the experimenter and intermixed into a continuous stream of non-target items (e.g. Farwell & Donchin, 1991; Rosenfeld et al., 1988; Rosenfeld et al., 2008), some of which consist of crime-relevant information (so-called probes). Probe items will stand out against the background of other irrelevant non-targets for a ‘guilty’ person only. Critically, the rare meaningful probes elicit a larger P300 ERP component relative to irrelevant non-targets. This procedure should therefore enable the examiner to detect crime-related knowledge in a guilty participant (but see Rosenfeld, Soskins, Bosh, & Ryan, 2004).

While most ERP studies using CITs have tested verbal material, some previous experiments have also examined concealed face recognition. In these experiments, participants are typically asked to make explicit familiarity judgments to familiar and unfamiliar faces. Genuinely familiar faces categorised as unfamiliar by the participants in an attempt to conceal familiarity (i.e., probes) elicit a larger P300 relative to irrelevant unfamiliar control faces (Meijer, Smulders, Merckelbach, & Wolf, 2007; Meijer, Smulders, & Wolf, 2009). Importantly, these studies have either used a single picture or very limited sets of images for each of the presented identities. However, a suspect might have seen images of the critical face in the media or during previous interviews, and recognising these particular images in a face CIT will accordingly not necessarily imply guilty knowledge. If researchers are aware of this situation, they might think that a test is useless, as a positive CIT using the same or similar images as those seen before will be difficult to interpret. In contrast, a test based on multiple ambient images should still be informative. If the probe identity is truly unfamiliar it will be unlikely to be recognised from highly variable images even if a specific picture is known, which will correctly result in a negative test. If the probe identity is truly familiar, it will be recognised from nearly all images, which in turn should result in a correct positive test result. It therefore appears desirable to establish a largely automatic face recognition response that is elicited by a wide range of images showing a critical identity.

1.3 Summary of Procedure

The present experiments introduced three modifications to the paradigm used in our previous experiments (Wiese, Tüttenberg, et al., 2019; see Table 1). In Experiment 1, we examined the effects of face familiarity while passively viewing a sequence of photos, asking participants only to respond to infrequent butterfly stimuli. We used multiple highly variable images of one familiar face, presenting each specific image only once, and contrasted this condition with the repeated presentation of a single image of a second highly familiar person. By investigating the N170, N250 and SFE as well as a classifier using supervised learning (Bishop, 2006), we tested whether face and person recognition processes are similarly elicited by highly variable versus identical exemplars of a face. Moreover, we examined whether the highly variable and single image conditions would result in similar detection rates in individual participants using a bootstrapping technique (Di Nocera & Ferlazzo, 2000).

Table 1. Overview of experimental procedures.

	Experiment 1	Experiment 2	Experiment 3
Short Description	Incidental Face Recognition	Explicit Familiarity Judgments	Concealed Familiarity
Research Question	Does image repetition affect	Does an explicit familiarity task	How reliably can we detect fam-

	ERP measures of recognition?	enhance ERP measures of recog- nition?	iliarity when participants actively deny it?
Conditions	High Variability - Familiar High Variability - Unfamiliar Single Image - Familiar Single Image - Unfamiliar	Familiar Unfamiliar Critical Unfamiliar	Acknowledged Familiar Concealed Familiar Unfamiliar
Stimuli	Each participant presented with 50 images of familiar ID1 (once) 1 image of familiar ID2 (50 times) 50 images of unfamiliar ID1 (once) 1 image of unfamiliar ID2 (50 times) 8 images of butterflies	40 images of each of 22 IDs, 14 female, 8 male Each participant presented with 1 familiar ID 2 unfamiliar IDs (one "critical")	40 images of each of 27 IDs, 20 female, 7 male Each participant presented with familiar ID1 familiar ID2 unfamiliar ID
Task	"Press key whenever a butter- fly is presented!"	"Press [left/right] key for familiar, [right/left] for unfam- iliar faces!"	"Press [left/right] key for familiar ID1 ('Acknowledge!'), press [right/left] for familiar ID2 ('Conceal!') and unfamiliar ID!"
N	22	19	19

In Experiment 2, we changed the task to an explicit familiarity decision. Participants were presented with multiple ambient images of two unfamiliar faces and one familiar face and were instructed to respond according to their true familiarity with these identities. Analysing the N250 effect and SFE across experiments enabled us to test a potential modulation of face recognition by explicit, intentional identity processing. Due to the use of two unfamiliar identities, the experiment additionally enabled us to estimate a “false alarm rate”, i.e. how likely a “critical” truly unfamiliar face would be wrongly classified as familiar in individual participants’ EEG responses.

Experiment 3 further examined voluntary control by directly comparing acknowledged and concealed familiar face recognition. We asked participants to correctly indicate their familiarity with one of two familiar identities and to correctly deny knowing a truly unfamiliar identity. Critically, participants were further instructed to “lie” about a second truly familiar face, i.e. to respond “unfamiliar” whenever this identity was presented.

This allowed us to contrast the intention to recognise a face with the intention to hide familiarity. By using a similar structure to previous CIT experiments, we were additionally able to estimate a "hit rate", i.e. how well the bootstrapping procedure would detect an attempt to conceal familiarity with the critical identity in individual participants.

2. Experiment 1: Incidental recognition

Experiment 1 tested familiar face recognition under incidental viewing conditions and additionally examined the role of image variability for familiar face recognition. It is known that within-person image variability is critical for learning a new facial identity (Burton, Kramer, Ritchie, & Jenkins, 2016; Jenkins & Burton, 2011; Ritchie & Burton, 2017). To recognise a face from a picture that has not been seen before, participants need substantial experience with exactly how different a specific person can look (Kramer, Young, & Burton, 2018). However, image variability might also be a critical aspect of familiar face recognition. To continuously trigger the processing of individual identity rather than of specific images, it might be necessary to present highly variable images of a known person (for a related approach using release from adaptation with MRI, see Davies-Thompson, Gouws, & Andrews, 2009; Weibert et al., 2016).

Using a butterfly detection task, participants were presented with multiple highly variable or single images of personally familiar and unfamiliar faces. First, we expected to replicate our previous findings of substantial ERP familiarity effects using ambient images of highly familiar faces (Wiese, Tüttenberg, et al., 2019). Moreover, as previous face-CIT studies have used a single picture or very limited sets of images (see above), we tested whether reduced facial identity processing would be observed when the same image of a person is repeatedly presented (see above). Assuming adaptation to the specific picture in the single image condition (Davies-Thompson et al., 2009; Weibert et al., 2016), we predicted reduced N250 familiarity effects and SFEs in this relative to the high variability condition. Finally, we predicted substantially smaller proportions of participants with reliable familiarity effects at the individual subject level in the single image relative to the high variability condition.

2.1 Methods

2.1.1 Participants

We tested 24 Durham University undergraduate students, two of whom were excluded due to technical problems during EEG recording. The final sample of 22 participants consisted of 19 females and 3 males, with a mean age of 20.1 years (SD = 0.9). All participants were right-handed according to a modified version of the Edinburgh Handedness Inventory (Oldfield, 1971) and had normal or corrected to normal vision. None suffered from neurological or psychiatric disorders or took central-acting medication. Participants received either £7.50/h or course credit for compensation. All gave written informed consent, and the study was approved by the ethics committee of Durham University's Psychology Department.

2.1.2 Stimuli

Participants were asked to provide the experimenters with images of two highly personally familiar faces (close friends, relatives etc.) known from outside the university. While 50 images of the first identity were required (high variability condition), only one image of the other identity was collected (single image condition). It was emphasised, that both identities needed to be similarly highly familiar, and level of familiarity was checked after the main experiment (see below). All depicted individuals gave written informed consent to the use of their pictures for the purposes of the experiment. Eight different pictures of butterflies were used as targets. Participants were tested in pairs, and images were balanced across conditions within each pair, such that the two familiar identities of the first participant were used as the unfamiliar identities for the second participant, and vice versa.

Rectangles around the faces and butterflies were cropped from the original images, re-sized to 190 x 285 pixels, and converted to greyscale. All images were adjusted for luminance using the SHINE toolbox (Willenbockel et al., 2010). During the experiment, stimuli were presented in the centre of the screen against a grey background.

2.1.3 Experimental Design and Procedure

Participants were seated in an electrically shielded and sound-attenuated cabin with their head in a chin rest at a distance of 80 cm from a computer screen. During the experiment, participants saw the 50 different images of the first familiar and unfamiliar identities, 50 repetitions of the same picture of the other familiar and unfamiliar identities, respectively, and 20 trials with pictures of butterflies in random order. Each trial started with a fixation cross which varied randomly in duration from 1,500 ms to 2,500 ms, followed by a face stimulus presented for 1,000 ms. Participants were only informed that they would see pictures of familiar faces, unfamiliar faces, and butterflies, and asked to press a response button with their right index finger as fast as possible whenever a picture of a butterfly was presented.

After the main experiment, participants were asked to rate each identity used in their respective version of the experiment for familiarity ("How likely would you recognise this person?", from 1 = very unlikely to 5 = very likely), valence ("How do you feel when you see this person?", from 1 = very positive to 5 = very negative), and arousal ("How do feel when you see this person?", from 1 = very excited to 5 = not excited at all). Valence and arousal ratings were illustrated using the Self-Assessment Mannequin scale (Bradley & Lang, 1994). For the two identities in the high variability condition, eight randomly selected images were shown simultaneously on the screen during this task. For the two identities in the single image condition, only the one available picture was presented.

2.1.4 EEG recording and data analysis

64-channel EEG (EEGo, ANT Neuro, Hengelo, The Netherlands) was recorded from DC to 200 Hz with a sample frequency of 1024 Hz using sintered Ag/Ag-Cl electrodes mounted in a textile cap. An electrode on the forehead (AFz) served as ground, and CPz was used as the recording reference. Recording sites corresponded to an extended 10-20 system, including ventral electrode positions such as TP9/TP10, P9/P10, and PO9/PO10.

Blink artefacts were corrected using BESA 6.0 (BESA GmbH, Graefelfing, Germany). EEG was then segmented from -200 to 1,000 ms relative to stimulus onset, with the first 200 ms as the baseline. Trials with non-ocular artefacts and saccades were rejected using the BESA 6.0 toolbox with an amplitude threshold of 100 µV and a gradient criterion of 75 µV. Remaining trials were re-calculated to the common average reference, digitally low-pass filtered at 40 Hz (12 dB/oct, zero-phase shift), and averaged according to experimental conditions. Only trials with correct responses were analysed. The average number of trials was 46.2 (SD = 3.7, min = 33) in the high variability familiar condition, 46.1 (SD = 3.9, min = 38) in the high variability unfamiliar condition, 46.1 (SD = 3.5, min = 38) in the single image familiar condition, and 45.6 (SD = 3.2, min = 39) in the single image unfamiliar condition.

In the resulting waveforms, mean amplitudes for the N170 (140-180 ms), N250 (200 – 400 ms) and the SFE (400-600 ms) were calculated at electrodes TP9/TP10 and P9/P10. Time windows and electrodes were chosen prior to the analysis on the basis of our previous studies (Wiese, Ingram, et al., 2019; Wiese, Tüttenberg, et al., 2019), and they are consistent with the timing and scalp distribution of familiarity effects observed in the present study (see Figure 1). Statistical analyses of ERP data were carried out using Analyses of Variance (ANOVA), with the within-subjects factors hemisphere (left, right), site (TP, P), familiarity (familiar, unfamiliar), and variability (high variability, single image), as well as t-tests for planned comparisons. Following an estimation approach in data analysis (Cumming, 2012), we report measures of effect size with appropriately sized confidence intervals (CIs) throughout. CIs for partial eta squared were calculated using scripts provided by M.J. Smithson (www.michaelsmithson.online/stats/CIstuff/CI.html). Cohen's d for repeated-measures t-tests was bias-corrected and calculated using the mean standard deviation rather than the standard deviation of the difference as the denominator ($d_{unb.}$) using ESCI (Cumming, 2012).

The reliability of familiarity effects in individual participants was tested using bootstrapping (Di Nocera & Ferlazzo, 2000). For both the N250 and SFE time window, individual participants' EEG epochs were randomly re-assigned to familiar versus unfamiliar conditions, for high variability and single image conditions, respectively. This procedure was carried out 10,000 times, and differences between re-assigned conditions were calculated. Reliable effects were assumed if the actual individual familiarity effect at TP9/TP10 was larger than 95% of re-samplings (Wiese, Tüttenberg, et al., 2019).

Finally, we applied classification analysis using supervised learning (Bishop, 2006). These analyses were carried out using the MVPA-Light toolbox (Treder, 2020), modified to allow stochastic noise to be added to the training data. For each participant, a logistic regression was performed separating pairs of conditions (i.e., high variability familiar versus unfamiliar, single image familiar versus unfamiliar) as follows. For each channel, the average of the first 200 ms (pre-stimulus) was subtracted from the signal. Channels P9, P10, TP9, TP10 were downsampled to 1/100 (1/50 for experiment 3), converting 1024 samples at 1024 Hz (512 samples at 512 Hz for Experiment 3) into 11 timepoints separated by 97.7 ms. For the main analysis we only used these four channels as they showed the strongest familiarity effects during initial ERP analysis, although we later analysed the performance from all channels (see below). The classifier was therefore working with a 44-dimensional dataset (4 channels x 11 timepoints), with 40-50 trials per condition (varying across participants due to the artefact rejection procedure described above). Performance of the logistic regression classifier was assessed using ten-fold cross-validation across trials, which consisted of training the model on 90% of the trials and testing it on the remaining 10%. This was then repeated for the next 10% of the trials until all data was used for testing. This process itself was then repeated 10 times with shuffled trials in order to increase repeatability. In addition we added stochastic noise to the training data for regularisation (Bishop, 1995). This classification technique therefore provided a percentage of the number of trials that were correctly classified for each participant, for each pair of familiarity conditions.

In addition, we examined the classification performance across time, based on the same electrodes and parameters as above (but assessed for each of the 11 timepoints separately). Finally, to assess the spatial extent of the classification performance we performed a Searchlight analysis based on all channels (Kriegeskorte, Goebel, & Bandettini, 2006), providing maps of the classification performance across channels. Data, analysis code and materials (except for stimuli, as we do not have permission to publish facial photographs) have been archived in a publicly accessible repository (https://osf.io/7xtdy/?view_only=699019a1ec6249a6881067f9449e79cc).

2.2 Results

2.2.1 Performance

Participants performed close to ceiling in the butterfly detection task; hit rate: M = .98, SD = .03; false alarm rate: M = .003, SD = .005. Mean response time was 520 ms, SD = 62.

2.2.2 Event-related potentials

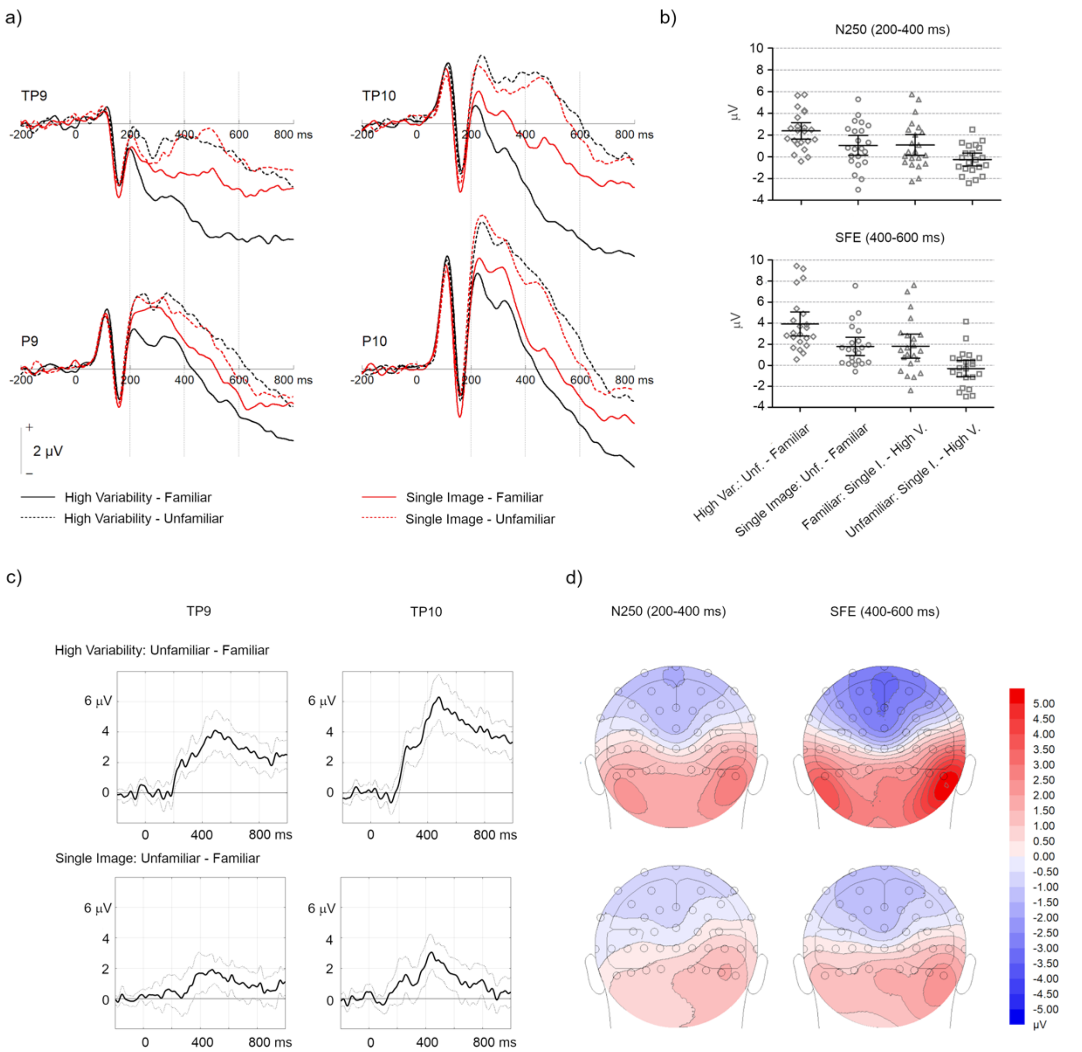

Figure 1. a) Grand average event-related potentials at left- and right-hemispheric electrodes TP9/TP10 and P9/P10. Dashed lines mark the N250 (200-400 ms) and SFE (400-600 ms) time ranges. b) Mean (+/- 95% CI) and individual familiarity effects in the N250 and SFE time ranges at electrodes TP9/TP10/P9/P10. c) Mean (+/- 95% CI) difference waves at left and right occipito-temporal electrodes TP9/TP10. d) Scalp-topographical voltage maps (spherical spline interpolation, 110 degrees equidistant projection) of familiarity effects in the N250 and SFE time window.

Visual inspection of ERP data suggested clear familiarity effects starting approximately 200 ms after stimulus onset. Importantly, familiarity effects were substantially larger in the high variability relative to the single image condition. These observations were confirmed by statistical analyses. A repeated-measures ANOVA in the N170 time range (140-180 ms) yielded trends for a main effect of variability, $F(1, 21) = 3.48$, $p = .076$, $\eta_p^2 = .14$, 90% CI [0, .36], with slightly larger amplitudes in the single image condition, as well as for an interaction of site by hemisphere by familiarity, $F(1, 21) = 4.34$, $p = .050$, $\eta_p^2 = .17$, 90% CI [0, .38]. However, separate follow-up tests at each of the four electrodes did not reveal any significant familiarity effects, all $F < 1$, all $p > .72$, all $\eta_p^2 < .05$.

A corresponding analysis in the N250 time range (200-400 ms) revealed a significant main effect of familiarity, $F(1, 21) = 29.67$, $p < .001$, $\eta_p^2 = .59$, 90% CI [.32, .71], as well as a significant interaction of familiarity by hemisphere, $F(1, 21) = 6.53$, $p = .018$, $\eta_p^2 = .24$, 90% CI [.02, .45], reflecting larger familiarity effects over the right hemisphere (see Figure 1d). Importantly, the interaction of familiarity by variability was significant, $F(1, 21) = 6.90$, $p = .016$, $\eta_p^2 = .25$, 90% CI [.03, .45]. Planned comparisons revealed more negative-going N250 amplitudes in the high variability familiar, M = 0.10 µV, 95% CI [-1.39, 1.59], relative to the high variability unfamiliar condition, M = 2.48 µV, 95% CI [0.95, 4.01], $t(21) = 6.50$, $p < .001$, $d_{unb.} = 0.70$, 95% CI [0.40, 0.99]. Similarly, for single images, N250 was more negative-going for the familiar, M = 1.19 µV, 95% CI [-0.29, 2.66], relative to the unfamiliar condition, M = 2.24 µV, 95% CI [0.88, 3.60], $t(21) = 2.40$, $p = .026$, $d_{unb.} = 0.32$, 95% CI [0.04, 0.61], although the effect was considerably smaller than in the high variability condition. While familiar faces were significantly more negative in the high variability relative to the single image condition, $t(21) = 2.36$, $p = .028$, $d_{unb.} = 0.31$, 95% CI [0.04, 0.61], the corresponding comparison was not significant for unfamiliar faces, $t(21) = -0.88$, $p = .388$, $d_{unb.} = -0.07$, 95% CI [-0.24, 0.09].

A corresponding ANOVA in the SFE time range (400-600 ms) revealed a significant main effect of familiarity, $F(1, 21) = 44.42$, $p < .001$, $\eta_p^2 = .68$, 90% CI [.44, .78], qualified by interactions of familiarity by hemisphere, $F(1, 21) = 10.73$, $p = .004$, $\eta_p^2 = .34$, 90% CI [.08, .53], and familiarity by site, $F(1, 21) = 20.16$, $p < .001$, $\eta_p^2 = .49$, 90% CI [.21, .64], reflecting larger familiarity effects over the right hemisphere and at more anterior sites (see Figure 1d). A significant interaction of familiarity by variability, $F(1, 21) = 19.73$, $p < .001$, $\eta_p^2 = .48$, 90% CI [.20, .64], was further qualified by an interaction of familiarity by variability by site, $F(1, 21) = 5.62$, $p = .027$, $\eta_p^2 = .21$, 90% CI [.01, .42], suggesting a stronger interaction of the two experimental factors at more anterior sites. Planned comparisons again revealed significantly more negative amplitudes for high variability familiar, M = -2.65 µV, 95% CI [-4.36, -0.94], relative to unfamiliar faces, M = 1.27 µV, 95% CI [-0.21, 2.75], $t(21) = 7.08$, $p < .001$, $d_{unb.} = 1.05$, 95% CI [0.64, 1.53]. Similarly, single image familiar faces elicited more negative-going amplitudes, M = -0.83 µV, 95% CI [-1.94, 0.28], relative to unfamiliar faces, M = 0.97 µV, 95% CI [0.06, 1.87], $t(21) = 4.29$, $p < .001$, $d_{unb.} = 0.76$, 95% CI [0.35, 1.21], although the effect was again substantially smaller. Familiar faces were significantly more negative in the high variability relative to the single image condition, $t(21) = 3.83$, $p = .001$, $d_{unb.} = 0.69$, 95% CI [0.28, 1.13], but unfamiliar faces were not, $t(21) = -0.82$, $p = .423$, $d_{unb.} = -0.11$, 95% CI [-0.37, 0.16].

Bootstrapping in the N250 time segment revealed that 17/22 participants showed reliable familiarity effects in the high variability condition, proportion (P) = .77, 95% CI [.57, .90], while 7/22 participants demonstrated reliable familiarity effects in the single image condition, P = .32, 95% CI [.16, .53]. In the SFE time range, 21/22 participants yielded reliable familiarity effects in the high variability condition, P = .96, 95% CI [.78, .99], but only 10/22 showed a reliable familiarity effect in the single image condition, P = .46, 95% CI [.27, .65].

These results were complemented by classification analysis (see Figure 2a), which showed clearly above-chance (.5) correct classification of familiar versus unfamiliar face trials in both the high variability, M = .68, 95% CI [.65, .71], $t(21) = 12.23$, $p < .001$, $d_{unb.}$ = 2.51, 95% CI [1.67, 3.36][1], and single image condition, M = .63, 95% CI [.60, .66], $t(21) = 9.82$, $p < .001$, $d_{unb.}$ = 2.02, 95% CI [1.30, 2.74]. Importantly, classification performance was significantly better for the high variability condition, M = .05, 95% CI [.02, .08], $t(21) = 3.13$, $p = .005$, $d_{unb.}$ = 0.68, 95% CI [0.21, 1.19]. The time course of classification performance showed maximum accuracy between 400 and 600 ms in the high variability and between 200 and 500 ms in the single image condition. Finally, searchlight analysis of classification performance across the scalp revealed clear (right) occipito-temporal and mid-central maxima, likely reflecting opposite ends of the same underlying dipole (see upper part of Figure 2).

[1] CIs for one-sample t-test d were calculated following Hedges & Olkin (1985).

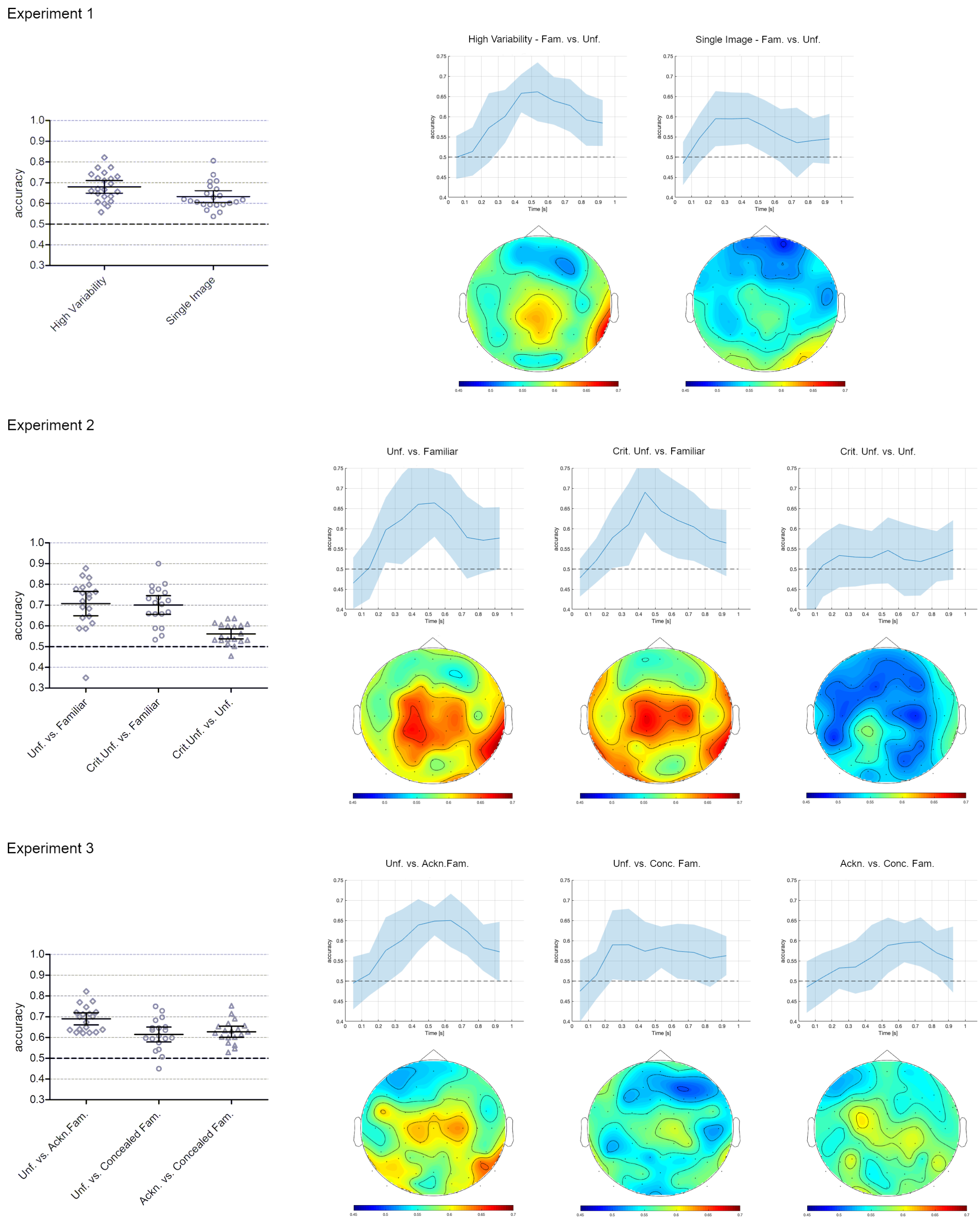

Figure 2. Classifier analyses for Experiments 1-3. Plots on the left show individual participant results (symbols) as well as means and 95% confidence intervals. Remaining columns show mean classifier performance (and 95% confidence intervals) at occipito-temporal electrodes TP9/TP10/P9/P10 over time (upper rows) and mean classifier performance across the scalp (lower rows).

2.2.3 Rating task

		High Variability		Single Image	
		Familiar	Unfamiliar	Familiar	Unfamiliar
Familiarity (1=unfamiliar, 5=highly familiar)	M	5.00	2.14	5.00	2.14
	SD	0.00	1.32	0.00	1.39
Valence (1=very positive, 5=very negative)	M	1.05	2.82	1.27	2.95
	SD	0.21	0.66	0.70	0.79
Arousal (1=highly arousing, 5=not arousing)	M	1.77	4.18	1.77	4.09
	SD	1.31	1.10	1.07	1.38

Table 2. Rating results from Experiment 1.

Rating results are reported in Table 2. Familiarity ratings revealed significant differences between familiar and unfamiliar faces both in the high variability condition, $t(21) = 10.18$, $p < .001$, $d_{unb.} = 2.96$[2], and in the single image condition, $t(21) = 9.66$, $p < .001$, $d_{unb.} = 2.81$. Both familiar faces were rated as maximally familiar, $t = 0$. Similarly, unfamiliar faces in the two variability conditions did not differ, $t = 0$.

Valence ratings yielded significantly more positive evaluations of familiar relative to unfamiliar faces, both in the high variability, $t(21) = 12.13$, $p < .001$, $d_{unb.} = 3.46$, and in the single image condition, $t(21) = 7.58$, $p < .001$, $d_{unb.} = 2.18$. Familiar faces were rated as highly positive and did not differ significantly between variability conditions, $t(21) = 2.02$, $p = .057$, $d_{unb.} = 0.42$, 95% CI [-0.01, 0.88]. Unfamiliar faces were rated as neutral and again did not differ, $t(21) = 0.72$, $p = .480$, $d_{unb.} = 0.18$, 95% CI [-0.33, 0.70].

Finally, participants rated familiar faces as significantly more arousing than unfamiliar faces, both in the high variability, $t(21) = 5.97$, $p < .001$, $d_{unb.} = 1.92$, 95% CI [1.10, 2.87], and in the single image condition, $t(21) = 6.10$, $p = .480$, $d_{unb.} = 1.81$, 95% CI [1.04, 2.70]. Neither familiar, $t = 0$, nor unfamiliar faces, $t(21) = -0.40$, $p = .693$, $d_{unb.} = -0.07$, 95% CI [-0.43, 0.29], were rated as differentially arousing in the two variability conditions.

[2] Note that ESCI only calculates CIs for estimates of δ between -2 and 2 (see Cumming, 2012, p. 306-307).

2.3 Discussion

Using incidental recognition, Experiment 1 directly tested whether image variability affects familiar face recognition by presenting multiple ambient images or single pictures of familiar and unfamiliar faces. We observed substantially smaller ERP familiarity effects, both in the N250 and SFE time windows, for single relative to highly variable images. Similarly, the classification analysis yielded more accurate separation of familiar from unfamiliar trials in the high variability relative to the single image condition. These findings suggests that face and person recognition processes are not engaged to the same extent in the former relative to the latter condition. It seems that single images are learnt during the experiment, and are then, at least partly, recognised based on specific pictorial rather than structural codes (Bruce & Young, 1986; Bruce & Young, 2012). Picture recognition presumably reduces the effort spent on the more costly and time-consuming matching of structural codes with image-independent representations. Therefore, engaging in picture rather than face recognition appears plausible, given that it is highly inefficient to process similar instances of the same face in depth over and over again. The trend for a larger N170 response for single images may be seen as supporting this suggestion, as it may reflect picture learning during the experiment (see Caharel et al., 2005, for a similar finding of larger N170 amplitudes for more often repeated images). Consequently, highly variable images of a given identity will continuously trigger familiar face recognition processes to their full extent, while the repeated presentation of a single image will not.

Similarly, a substantially higher proportion of individual participants demonstrated ERP familiarity effects in the high variability condition, with a proportion of .96 showing a reliable SFE in this relative to only .46 in the single image condition. Arguably, the use of multiple ambient images is therefore not only theoretically interesting but also clearly preferable in a potential applied scenario. In conclusion, Experiment 1 clearly supports the considerable importance of within-person variability for familiar face recognition and replicates our previous finding that familiarity with a known face can be established with high accuracy without explicitly asking the participant (Wiese, Tüttenberg, et al., 2019). Experiments 2 and 3 will turn to the question whether the intention to recognise or to conceal familiarity influences the ERP effects.

3. Experiment 2: Explicit Familiarity Judgments

Experiment 2 examined the role of intentionality for familiar face recognition. We presented multiple ambient images of two unfamiliar and one familiar identities and, as an important deviation from our previous experiments, asked our participants to explicitly categorise each picture as either showing a familiar or unfamiliar face. Any change in ERP familiarity effects, relative to our previous experiments, would therefore likely reflect the influence of intentional rather than incidental processing of familiarity. As the N250 seems to be largely unaffected by task demands or conscious awareness (Eimer et al., 2012; Wiese, Ingram, et al., 2019), we predicted highly similar familiarity effects relative to our previous studies. By contrast, as the SFE depends more strongly on the attentional resources available for face processing, we predicted larger familiarity effects in this time window relative to Experiment 1.

The experiment resembled the basic paradigm of previous face CIT studies by using target ("familiar"), probe ("critical unfamiliar"), and irrelevant ("unfamiliar control") conditions (Meijer et al., 2007; Meijer et al., 2009). Importantly, however, we used multiple images for each facial identity. The experiment therefore aimed to establish baseline ERP familiarity effects in an explicit recognition task. Moreover, the inclusion of two unfamiliar faces allowed us to examine whether individual participants would demonstrate reliable differences between these identities, and therefore to determine a “false alarm rate” in a CIT-like paradigm.

3.1 Methods

3.1.1 Participants

We tested 19 undergraduate students from Durham University (11 female, 8 male; mean age = 21.1 years, SD = 1.0). Exclusion criteria and compensation were identical to Experiment 1. All participants gave written informed consent, and the experiment was approved by the ethics committee of Durham University’s Department of Psychology.

3.1.2 Stimuli, Experimental Design and Procedures, Data Analysis

Across the experiment, stimuli consisted of 40 different images of each of 22 different identities (14 female, 8 male). Each participant was presented with pictures of three identities, one of which was highly personally familiar while the other two were unfamiliar. Unfamiliar identities were randomly assigned to ‘critical unfamiliar’ and ‘unfamiliar’ control conditions. Stimuli in each triplet of identities were chosen such that those requiring the same key press were never of the same gender to minimise response strategies unrelated to face identity. As with Experiment 1, all face stimuli were provided by asking participants to bring multiple photos of a highly familiar person – with these being used as unfamiliar stimuli for different participants.

All experimental procedures were identical to Experiment 1, except that participants were informed that they would see pictures of one familiar face and two unfamiliar faces, and were asked to correctly indicate their familiarity with all presented identities via left and right index finger button presses. Key assignment was balanced across participants. EEG data analysis was analogous to Experiment 1. Mean number of trials were 36.8 (SD = 1.9, min = 33) for the familiar condition, 37.5 (SD = 2.2, min = 30) for the critical unfamiliar condition, and 36.2 (SD = 4.0, min = 25) for the unfamiliar control condition.

3.2 Results

3.2.1 Performance

Mean correct response times for familiar (M = 553 ms, SD = 63), critical unfamiliar (M = 536 ms, SD = 80), and unfamiliar control conditions (M = 538 ms, SD = 74) did not differ, $F(2, 36) = 1.22$, $p = .307$, $\eta_p^2 = .06$, 90% CI [0, .19]. Similarly, accuracies for familiar (M = .95, SD = .03), critical unfamiliar (M = .94, SD = .16), and unfamiliar control conditions (M = .97, SD = .03) were not significantly different, $F(2, 36) = 0.56$, $p = .578$, $\eta_p^2 = .03$, 90% CI [0, .13].

3.2.2 Event-related potentials

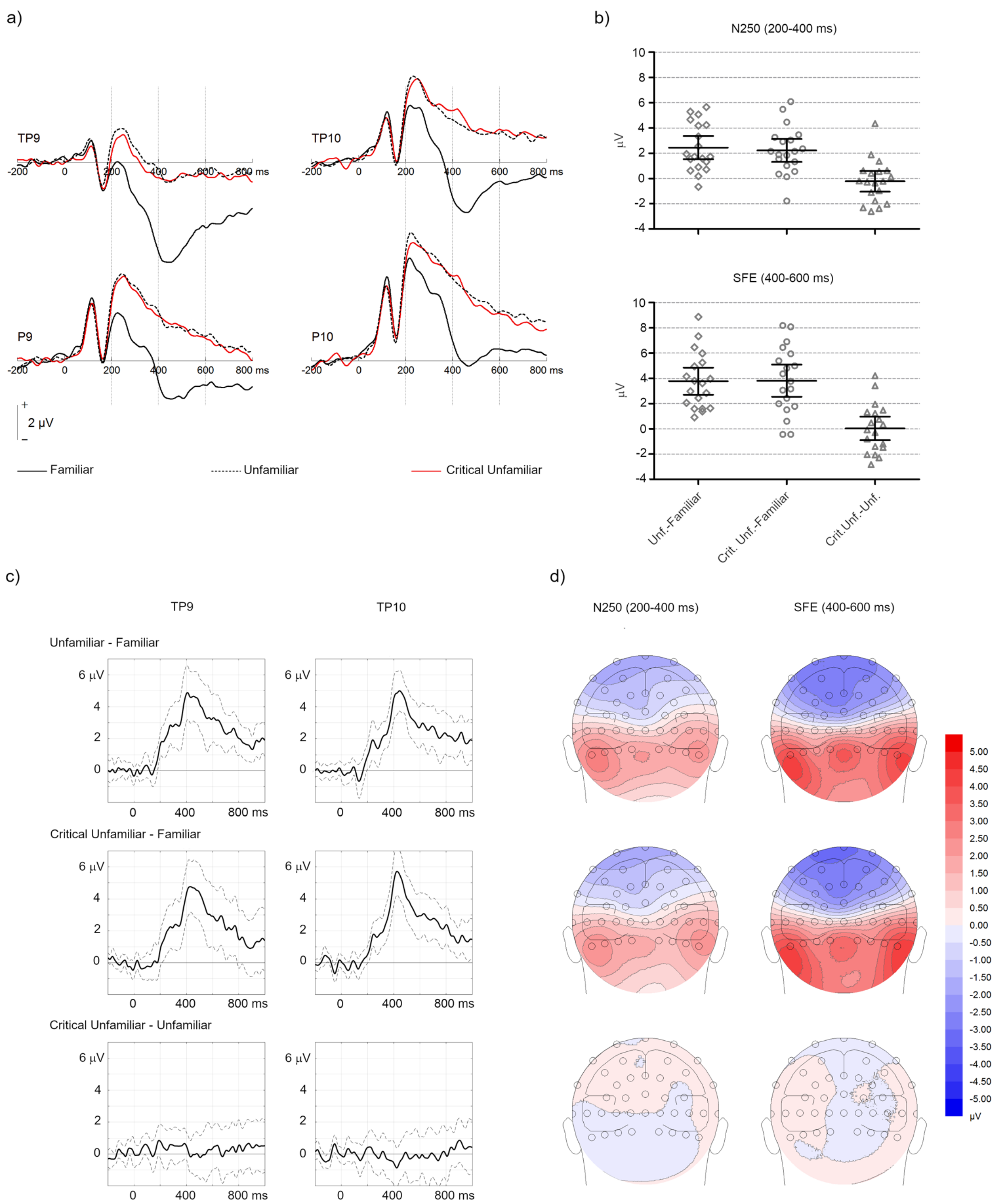

Figure 3. a) Grand average event-related potentials from Experiment 2 at left- and right-hemispheric electrodes TP9/TP10 and P9/P10. Dashed lines mark the N250 (200-400 ms) and SFE (400-600 ms) time ranges. b) Mean (+/- 95% CI) and individual familiarity effects in the N250 and SFE time ranges at electrodes TP9/TP10/P9/P10. c) Mean (+/- 95% CI) difference waves at left and right occipito-temporal electrodes TP9/TP10. d) Scalp-topographical voltage maps (spherical spline interpolation, 110 degrees equidistant projection) of familiarity effects in the N250 and SFE time window.

Visual inspection of ERP data suggested clear familiarity effects, for both familiar versus critical unfamiliar and familiar versus unfamiliar control conditions, starting approximately 200 ms after stimulus onset (see Figure 3). These observations were confirmed in statistical analyses. A repeated-measures ANOVA in the N170 time range (140-180 ms) with the within-subject factors hemisphere (left, right), site (TP, P) and familiarity (familiar, critical unfamiliar, unfamiliar control) did not reveal any significant effects involving the familiarity factor, all $F < 1.20$, all $p > .30$, all $\eta_p^2 = .07$. By contrast, a corresponding ANOVA in the N250 time range (200-400 ms) revealed a significant main effect of familiarity, $F(2, 36) = 20.76$, $p < .001$, $\eta_p^2 = .54$, 90% CI [.32, .64]. Planned t-tests yielded significantly more negative-going amplitudes in the familiar, M = 1.25 µV, 95% CI [-0.03, 2.53], relative to the unfamiliar control condition, M = 3.70 µV, 95% CI [1.84, 5.56], $t(18) = 5.63$, $p < .001$, $d_{unb.} = 0.71$, 95% CI [0.38, 1.09]. Similarly, the familiar condition was more negative than the critical unfamiliar condition, M = 3.48 µV, 95% CI [1.93, 5.02], $t(18) = 5.15$, $p < .001$, $d_{unb.} = 0.72$, 95% CI [0.37, 1.13]. Critical unfamiliar and unfamiliar control conditions did not differ, $t(18) = -0.58$, $p = .569$, $d_{unb.} = -0.06$, 95% CI [-0.28, 0.15].

A corresponding analysis of the SFE time window (400-600 ms) again yielded a significant main effect of familiarity, $F(2, 36) = 35.01$, $p < .001$, $\eta_p^2 = .66$, 90% CI [.48, .74]. Planned t-tests yielded significantly more negative-going amplitudes in the familiar, M = -2.15 µV, 95% CI [-3.85, -0.45], relative to the unfamiliar control condition, M = 1.63 µV, 95% CI [-0.23, 3.49], $t(18) = 7.41$, $p < .001$, $d_{unb.} = 0.98$, 95% CI [0.59, 1.44]. Similarly, the familiar condition was more negative than the critical unfamiliar condition, M = 1.67 µV, 95% CI [0.15, 3.18], $t(18) = 6.29$, $p < .001$, $d_{unb.} = 1.09$, 95% CI [0.62, 1.65]. Critical unfamiliar and unfamiliar control conditions did not differ, $t(18) = 0.09$, $p = .933$, $d_{unb.} = 0.01$, 95% CI [-0.24, 0.26].

To more directly test the role of explicit recognition, we compared ERP familiarity effects from Experiment 1 (high variability unfamiliar – familiar) to those from Experiment 2 (unfamiliar control – familiar). An independent t-test in the N250 time range revealed no significant difference between experiments, $t(39) = -0.12$, $p = .908$, $d_{unb.} = -0.04$, 95% CI [-

0.65, 0.58]. The corresponding test for the SFE again yielded no significant difference, $t(39) = 0.18$, $p = .855$, $d_{unb.} = 0.06$, 95% CI [-0.56, 0.67].

Bootstrapping in the N250 time range revealed reliable familiarity effects in 13/19 participants for the familiar versus unfamiliar control condition, P = .68, 95% CI [.46, .85], in 12/19 participants for the familiar versus critical unfamiliar condition, P = .63, 95% CI [.41, .81], and in 5/19 participants for the critical unfamiliar versus unfamiliar condition, P = .26, 95% CI [.12, .49]. Corresponding analyses in the SFE time range yielded reliable familiarity effects in 14/19 participants for the familiar versus unfamiliar control condition, P = .74, 95% CI [.51, .88], and in 15/19 participants for the familiar versus critical unfamiliar condition, P = .79, 95% CI [.57, .92]. 3/19 participants demonstrated false positive familiarity effects when comparing the critical unfamiliar with the unfamiliar control condition, P = .16, 95% CI [.06, .38].

Classifier analysis (see Figure 2b) showed clearly above-chance level separation of familiar from both unfamiliar, M = .71, 95% CI [.65, .77], $t(18) = 7.43$, $p < .001$, $d_{unb.} = 1.63$, 95% CI [0.96, 2.31], and critical unfamiliar face trials, M = .70, 95% CI [.66, .75], $t(18) = 9.35$, $p < .001$, $d_{unb.} = 2.05$, 95% CI [1.27, 2.84]. The classifier also correctly separated unfamiliar and critical unfamiliar faces above chance-level, M = .56, 95% CI [.54, .59], $t(18) = 5.35$, $p < .001$, $d_{unb.} = 1.18$, 95% CI [0.61, 1.75]. Moreover, classifier performance was more accurate for familiar versus unfamiliar faces than for critical unfamiliar versus unfamiliar faces, $M_{diff.} = .15$, 95% CI [.09, .21], $t(18) = 5.09$, $p < .001$, $d_{unb.} = 1.51$, 95% CI [0.77, 2.35], as well as for familiar versus critical unfamiliar than critical unfamiliar versus unfamiliar face trials, $M_{diff.} = .14$, 95% CI [.09, .19], $t(18) = 5.75$, $p < .001$, $d_{unb.} = 1.78$, 95% CI [0.97, 2.72]. At the same time, no difference in classifier performance was detected for familiar versus unfamiliar as compared to familiar versus critical unfamiliar face trials, $M_{diff.} = .01$, 95% CI [-.04, .05], $t(18) = 0.35$, $p = .734$, $d_{unb.} = 0.06$, 95% CI [-0.30, 0.43]. Time course analysis revealed maximum classification performance between 400 and 600 ms, while spatial analysis again demonstrated (right) occipito-temporal and central maxima (see middle part of Figure 2).

3.3 Discussion

Experiment 2 examined the role of explicit face recognition by making familiarity task-relevant. Similar to our previous experiments using butterfly detection tasks, we observed clear ERP familiarity effects, both in the N250 and the SFE time window. Importantly, these effects were highly similar to those observed in Experiment 1 (with Cohen's d for the between-experiment comparison ranging between -0.04 and 0.06), and a comparison between experiments did not result in any significant differences. These findings suggest that the examined ERP effects are not modulated by whether participants are explicitly responding to the familiarity of a face or not.

Together with previous findings indicating independence of awareness and no interference from additional tasks (Eimer et al., 2012; Wiese, Ingram, et al., 2019), and given its relatively fast occurrence after stimulus onset, the present results suggest that the N250 effect is indeed an automatic response according to the criteria outlined above. At the same time, even if the SFE is not modulated by voluntary control, it depends on the availability of processing resources (Wiese, Ingram, et al., 2019) and therefore does not reflect a fully automatic response. It appears, however, that it is elicited as long as participants have sufficient resources to attend to the face stimuli.

These findings were complemented by the classifier analysis which demonstrated discrimination of familiar and unfamiliar faces substantially above-chance level. Again, results are comparable to those obtained in Experiment 1, which suggests brain responses independent of task demands. Of note, the classifier also separated unfamiliar and critical unfamiliar faces above chance level, although with clearly reduced accuracy. This effect might reflect learning, as participants likely started to recognise some of the presented unfamiliar pictures as showing the same identity in the course of experiment.

Observations of similar familiarity effects in the two experiments at the group level, as reflected in large and similar effect sizes for the SFE, were only partly confirmed at the individual participant level. Most notably, whereas bootstrapping results in Experiment 1 suggested near perfect hit rates in the SFE time range in the high variability condition, a somewhat smaller proportion of participants (around 75%) demonstrated reliable effects in Experiment 2. It thus appears that the additional task demand of the familiarity judgment actually reduced the reliability of the signal rather than boosting it. As averaged ERPs were comparable between experiments, this reduction is most likely explained by larger variability in single-trial EEG responses in Experiment 2. This interpretation is also in line with the finding of a .16 false alarm rate when comparing the two unfamiliar faces. Again, this proportion is substantially higher than false alarm rates observed using a butterfly detection task (< .05; see Experiment 2 in Wiese, Tüttenberg, et al., 2019). It therefore appears that implicit tasks that do not require additional decisional and motor processes are preferable to create a reliable measure of familiar face recognition.

The conclusions we have drawn so far concerning participants' attempts to conceal their recognition are based on comparisons between experiments. In the final study we report the effects of manipulating the intention to recognise directly within a single experiment, asking participants sometimes to report their recognition of a familiar face and sometimes to conceal it.

4. Experiment 3: Concealed Familiarity

In Experiment 3, participants were presented with two familiar and one unfamiliar facial identity and were asked to make familiarity judgments. Crucially, however, while honestly responding to one familiar identity, they were asked to “lie” about the other one, and to judge its respective images as unfamiliar. This experiment therefore directly compared the intention to recognise one identity with the intention to hide recognition of the other familiar identity. If the N250 or SFE were modulated by voluntary control, we would assume larger effects for the acknowledged relative to the concealed face. In addition to these comparisons at the group level, the experiment also allowed tests of individual EEG responses when participants actively deny recognising a truly familiar face, a condition that resembles the detection of “guilty knowledge” in CIT experiments.

4.1 Methods

4.1.1 Participants

We tested 20 Durham University undergraduate students, one of whom was excluded due to technical problems during EEG recording. The final sample consisted of 14 female and 5 male participants, with a mean age of 20.2 years (SD = 1.0). Exclusion and inclusion criteria, as well as ethics, were identical to the previous experiments.

4.1.2 Stimuli

Stimuli consisted of 40 different ambient images of each of 27 different identities (20 female, 7 male). Each participant was presented with pictures of three identities, two of which were highly personally familiar. One familiar identity was a close friend known from university and images for that person were provided by the experimenters. The other familiar identity was a person known from outside the university (i.e., a close friend or relative), and images for this identity were provided by the participants. The third identity was unfamiliar, and images were taken from the personally familiar identities of a different participant. Editing of stimuli was analogous to the previous experiments.

4.1.3 Experimental design and procedure

Participants saw the 40 different pictures of each of the three identities in random order. Each trial started with a fixation cross which varied randomly in duration from 750 ms to 1,250 ms, followed by a face stimulus presented for 1,000 ms. The task consisted of familiarity decisions to the face stimuli via button presses using the left and right index fingers. Participants were informed that they would see pictures of two familiar faces and one unfamiliar face, and were instructed to correctly indicate that they are unfamiliar with the unfamiliar identity (unfamiliar condition) and familiar with one familiar identity chosen by the experimenter (acknowledged familiarity condition). Critically, participants were further asked to "lie" about the other familiar person, i.e., to incorrectly indicate that they are not familiar with the second familiar identity, and to conceal their familiarity with this identity (concealed familiarity condition). Assignment of familiar identities to the acknowledged versus concealed conditions, as well as key assignment was balanced across participants. Moreover, stimuli in each triplet of identities were chosen such that those requiring the same key press ("unfamiliar", either as a true response in the unfamiliar, or as a "lie" in the concealed familiarity condition) were never of the same gender to minimise response strategies unrelated to face identity.

4.1.4 EEG recording and data analysis

64-channel EEG (ANT Neuro, Hengelo, the Netherlands) was recorded from DC to 120 Hz with a sample frequency of 512 Hz. An electrode on the forehead (AFz) served as ground, and Cz was used as the recording reference. All other recording and analysis parameters were identical to the previous experiments. The average number of trials was 36.5 (SD = 2.8, min = 28) in the acknowledged familiarity condition, 35.8 (SD = 3.7, min = 25) in the concealed familiarity condition, and 36.4 (SD = 2.9, min = 30) in the unfamiliar condition.

4.2 Results

4.2.1 Performance

Correct response times for acknowledged familiarity (M = 612 ms, SD = 74), concealed familiarity (M = 607 ms, SD = 77), and unfamiliar faces (M = 592 ms, SD = 47) did not differ significantly, $F(2, 36) = 1.92$, $p = .161$, $\eta_p^2 = .10$, 90% CI [0, .23]. Similarly, accuracy for acknowledged familiarity (M = .95, SD = .03), concealed familiarity (M = .93, SD = .08), and unfamiliar faces (M = .95, SD = .04) was not significantly different, $F(2, 36) = 1.02$, $p = .370$, $\eta_p^2 = .05$, 90% CI [0, .17].

4.2.2 Event-related potentials

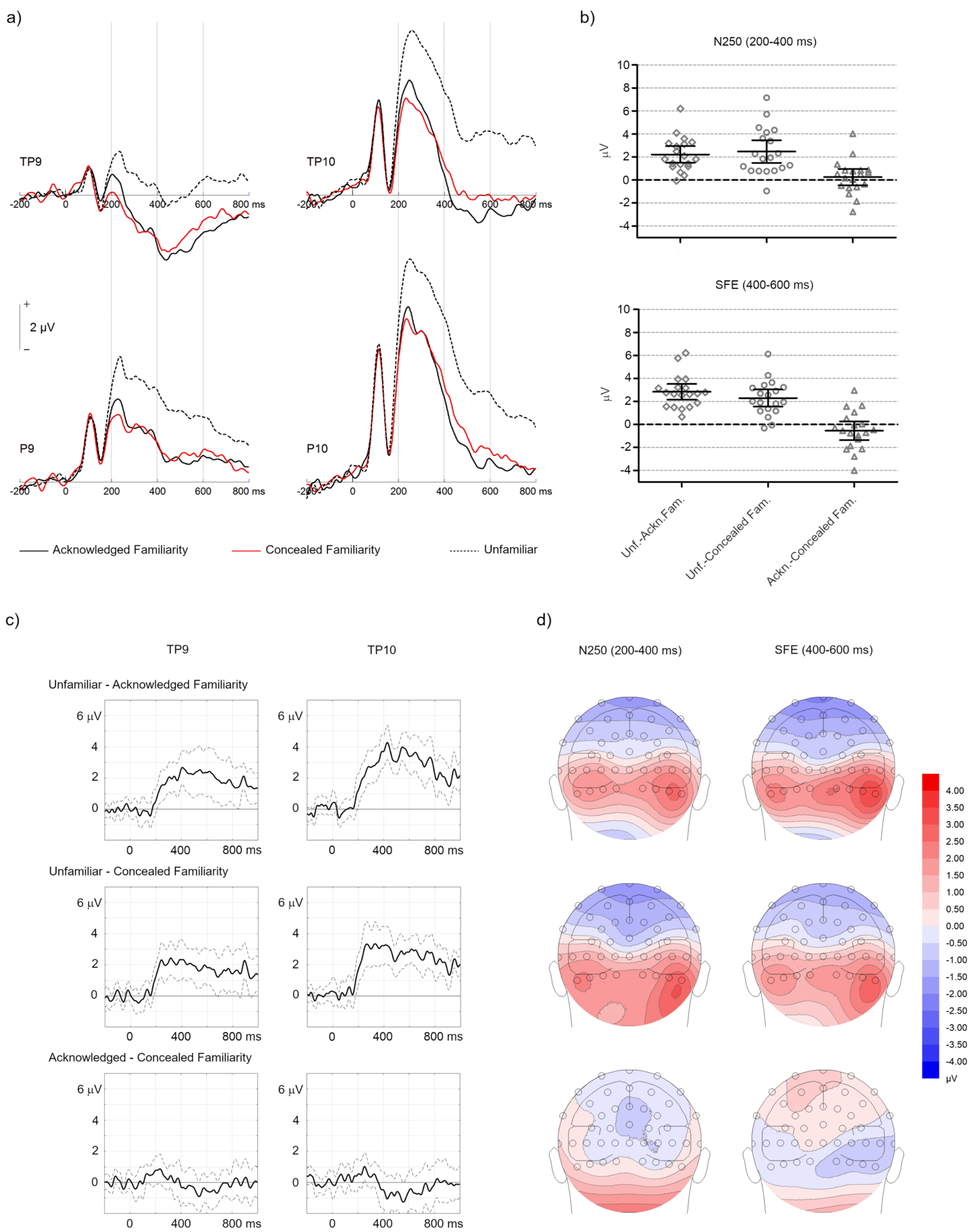

Figure 4. a) Grand average event-related potentials from Experiment 3 at left- and right-hemispheric electrodes TP9/TP10 and P9/P10. Dashed lines mark the N250 (200-400 ms) and SFE (400-600 ms) time ranges. b) Mean (+/- 95% CI) and individual familiarity effects in the N250 and SFE time ranges at electrodes TP9/TP10/P9/P10. c) Mean (+/- 95% CI) difference waves at left and right occipito-temporal electrodes TP9/TP10. d) Scalp-topographical voltage maps (spherical spline interpolation, 110 degrees equidistant projection) of familiarity effects in the N250 and SFE time window.

Visual inspection of ERP data again suggested clear familiarity effects starting approximately 200 ms after stimulus onset, both in the acknowledged and in the concealed familiarity conditions (see Figure 4c). This observation was confirmed by statistical analyses. A repeated-measures ANOVA in the N170 time range (140-180 ms) with the within-subjects factors hemisphere (left, right), site (TP, P), and familiarity (acknowledged familiarity, concealed familiarity, unfamiliar) did not result in any significant effects involving the familiarity factor, all $F < 1.13$, all $p > .333$, all $\eta_p^2 < .06$. By contrast, a repeated-measures ANOVA in the N250 time range (200-400 ms) revealed a significant main effect of familiarity, $F(2, 36) = 24.70$, $p < .001$, $\eta_p^2 = .58$, 90% CI [.37, .68]. Planned t-tests yielded significantly more negative-going amplitudes in the acknowledged familiarity, M = 2.95 µV, 95% CI [0.90, 4.99], relative to the unfamiliar condition, M = 5.17 µV, 95% CI [3.16, 7.17], $t(18) = 6.48$, $p < .001$, $d_{unb.} = 0.51$, 95% CI [0.29, 0.76]. Similarly, the concealed familiarity condition, M = 2.69 µV, 95% CI [0.78, 4.60], was more negative than the unfamiliar condition, $t(18) = 5.31$, $p < .001$, $d_{unb.} = 0.58$, 95% CI [0.31, 0.90]. Acknowledged and concealed familiarity conditions did not differ, $t(18) = 0.75$, $p = .463$, $d_{unb.} = 0.06$, 95% CI [-0.10, 0.23].

A corresponding ANOVA in the SFE time range (400-600 ms) again yielded a significant main effect of familiarity, $F(2, 36) = 35.89$, $p < .001$, $\eta_p^2 = .67$, 90% CI [.48, .75]. Planned t-tests yielded significantly more negative-going amplitudes in the acknowledged familiarity, M = -0.12 µV, 95% CI [-1.95, 1.71], relative to the unfamiliar condition, M = 2.72 µV, 95% CI [1.12, 4.31], $t(18) = 8.72$, $p < .001$, $d_{unb.} = 0.76$, 95% CI [0.48, 1.11]. Again, the concealed familiarity condition, M = 0.43 µV, 95% CI [-1.04, 1.89], was more negative than the unfamiliar condition, $t(18) = 6.47$, $p < .001$, $d_{unb.} = 0.69$, 95% CI [0.40, 1.04], while acknowledged and concealed familiarity conditions did not differ, $t(18) = -1.42$, $p = .174$, $d_{unb.} = -0.15$, 95% CI [-0.38, 0.07].

Bootstrapping in the N250 time window yielded reliable effects in 10/19 participants for the acknowledged familiarity versus unfamiliar condition, P = .53, 95% CI [.32, .73], in

13/19 participants for the concealed familiarity versus unfamiliar condition, P = .68, 95% CI [.46, .85], and in 1/19 participants for the acknowledged versus concealed familiarity condition, P = .05, 95% CI [.01, .25]. Corresponding analyses in the SFE time range revealed reliable effects in 14/19 participants for the acknowledged familiarity versus unfamiliar condition, P = .74, 95% CI [.51, .88], and in 10/19 participants for the concealed familiarity versus unfamiliar condition, P = .53, 95% CI [.32, .73]. Moreover, 3/19 participants, P = .16, 95% CI [0.06, .38], showed reliably stronger familiarity effects for the acknowledged relative to the concealed familiarity condition.

Finally, the classifier (see lower part of Figure 2) revealed above chance-level separation of both acknowledged versus unfamiliar, M = .69, 95% CI [.66, .72], $t(18) = 13.78$, $p < .001$, $d_{unb.} = 3.03$, 95% CI [1.97, 4.08], and concealed familiarity relative to unfamiliar face trials, M = .61, 95% CI [.58, .65], $t(18) = 6.68$, $p < .001$, $d_{unb.} = 1.47$, 95% CI [0.83, 2.10], as well as for acknowledged versus concealed familiarity trials, M = .63, 95% CI [.60, .65], $t(18) = 9.92$, $p < .001$, $d_{unb.} = 2.18$, 95% CI [1.36, 3.00]. Paired-sample t-tests yielded better performance for classifying acknowledged familiar versus unfamiliar than concealed familiar versus unfamiliar trials, $M_{diff.} = .08$, 95% CI [.03, .12], $t(18) = 3.51$, $p = .003$, $d_{unb.} = 1.07$, 95% CI [0.38, 1.82]. The classifier was also more accurate at separating acknowledged familiar versus unfamiliar trials than acknowledged versus concealed familiarity trials, $M_{diff.} = .06$, 95% CI [.03, .09], $t(18) = 4.13$, $p = .001$, $d_{unb.} = 1.03$, 95% CI [0.45, 1.69], but not at separating acknowledged versus concealed familiarity than concealed familiarity versus unfamiliar trials, $M_{diff.} = .01$, 95% CI [-.04, .07], $t(18) = 0.49$, $p = .628$, $d_{unb.} = 0.19$, 95% CI [-0.59, 0.98]. Analysis of the classifier time course revealed maximum accuracy between 500 and 700 ms in the acknowledged familiarity versus unfamiliar comparison, while maximum accuracy in the concealed familiar versus unfamiliar comparison was reached between 200 and 400 ms. Searchlight analysis again yielded occipito-temporal and central maxima of classifier performance (see Figure 2).

4.3 Discussion

Experiment 3 directly compared acknowledged and concealed recognition. We again observed clear familiarity effects both in the N250 and SFE time range, even when participants actively denied knowledge of a familiar face. In ERP group level analyses, acknowledged and concealed familiarity effects were indistinguishable. This finding confirms the conclusions proposed after the comparison of the first two experiments and suggests that the N250 familiarity effect is indeed largely automatic.

In addition, Experiment 3 examined how well an attempt to conceal knowledge of a truly familiar face would be detected at the individual participant level. We found reduced proportions of participants with reliable effects in the concealed relative to the acknowledged familiarity condition, most notably in the SFE time window, and about half of the participants were able to conceal their "guilty knowledge" from detection by the bootstrapping procedure. Similarly, the logistic regression-based classifier was more accurate at discriminating between acknowledged familiar versus unfamiliar faces than between concealed familiar versus unfamiliar faces. In combination with the results of Experiment 2, which found a false alarm rate of .16 when the critical face is truly unfamiliar using bootstrapping, as well as classifier performance with above-chance level discrimination between unfamiliar and critical unfamiliar faces, it seems that the EEG face CIT paradigm based on explicit judgments is moderately sensitive to detect true pre-experimental familiarity. This conclusion is further discussed below.

We note that ERP effects were somewhat less clear in Experiment 3 relative to both the other two experiments and to previous data (Wiese, Tüttenberg, et al., 2019). In particular, we did not observe a strong additional increase in the familiarity effect after the N250 time range (see Figure 4c), which was particularly evident in the concealed familiarity condition. We have observed a similarly reduced SFE before in conditions which direct attention away from the face stimulus (Wiese, Ingram, et al., 2019). It thus appears possible that denying true familiarity with one of the identities poses an additional task demand on the participants, for instance because they have to counter an initial tendency to respond "familiar". This in turn could distract participants from the face stimuli. At the same time, however, concealed and acknowledged familiarity did not elicit statistically significant different amplitudes in the SFE.

5. General Discussion

By varying task demands and analysing event-related potentials, the present series of experiments examined whether ERPs can serve as an index of a participant's recognition of a specific familiar face, and the extent to which this index might be useful even when the participant seeks to conceal their ability to recognise that face. Experiment 1 demonstrated substantial familiarity effects in both the N250 and SFE time windows when participants were undertaking an incidental task (butterfly detection). Furthermore, these effects were substantially increased when participants saw multiple images of the same person, compared to repetition of the same image. Experiments 2 and 3 then examined voluntary control (i) by changing the task from an implicit butterfly detection task to explicit familiarity judgments, and (ii) by directly comparing acknowledged and concealed familiarity. The results suggest that neither the N250 familiarity effect nor the SFE depend on the intention to recognise a face and add further evidence to the suggestion that the N250 effect is largely automatic.

The importance of image variability for unfamiliar face recognition and face learning is now widely accepted. Specifically, participants often find it difficult to match highly variable images of unfamiliar faces (Bruce et al., 1999; Jenkins et al., 2011) and profit from variability when learning new faces (Ritchie & Burton, 2017). While it is also well known that the recognition of familiar faces from different images is substantially easier (Burton, Wilson, Cowan, & Bruce, 1999; Jenkins et al., 2011), near perfect performance in matching or recognition tasks makes it difficult to examine the role of image variability for familiar face recognition with purely behavioural measures. The present study therefore examined ERP correlates of face recognition, namely the N250 familiarity effect and the SFE, and found enhanced effects on presentation of multiple photos for a viewed identity. Moreover, a classifier based on logistic regression was more accurate at separating familiar from unfamiliar faces in the high variability condition. High variability therefore seems to be important for triggering face recognition processes consistently during an experiment, presumably because participants learn to recognise specific pictures rather than matching a face with an abstract representation if the stimulus is repeatedly presented.

While Experiment 1 examined face familiarity under incidental conditions, Experiments 2 and 3 then examined the role of voluntary control by asking participants to make explicit familiarity responses. Our results show that the N250 effect does not depend on intentional processing as the effect was highly similar (i) for explicit familiarity judgment relative to the previously used implicit tasks, and (ii) for acknowledged relative to concealed familiarity. Given its relatively fast onset, and together with previous findings that have demonstrated the N250's independence of attentional load (Neumann & Schweinberger, 2008; Wiese, Ingram, et al., 2019), as well as its occurrence even in the absence of conscious awareness (Eimer et al., 2012), the present findings suggest that the effect is largely automatic. Interestingly, we also observed a similar SFE for concealed and acknowledged familiarity in the present study, although the effect occurs later in time than the N250 (and is therefore slower), and previous experiments have demonstrated that the effect depends on attentional resources (Wiese, Ingram, et al., 2019). Of note, while participants in these previous experiments had to direct attention away from the face stimuli to engage with the task at hand, faces were directly task relevant in the present Experiments 2 and 3. It therefore appears that the SFE is not modulated by the intention to acknowledge or conceal familiarity as long as participants are attending to the faces.

At the same time, effect sizes for the SFE were generally larger in Experiment 2 (with Cohen's d of 0.98 and 1.09 for the two relevant comparisons) relative to Experiment 3 (with d = 0.76 for acknowledged and d = 0.69 for concealed familiarity). Accordingly, it seems that the more complex task instruction in Experiment 3, which required the participants to acknowledge familiarity with one face but conceal familiarity with a different identity, led to additional cognitive load and in turn to overall smaller effects. This interpretation contrasts with previous studies on the P300 effect in face CITs, which have suggested that the effect at least partly reflects neural processes related to active lying (Meijer et al., 2009). As the SFE in the present study seems to get *smaller* with an instruction involving active lying, but is similar for concealed relative to the acknowledged recognition, the SFE seems to more directly reflect familiarity rather than lying.

The above considerations imply that the N250 and SFE reflect at least partially different processes. Alternatively, one might assume a single familiarity effect that persists for several hundred milliseconds. This latter interpretation, however, is not supported by the difference waves depicting familiarity effects. These waveforms typically show a sharp increase starting at approximately 200 ms. Importantly, the effect increases well beyond 400 ms, with a clear dip between 300 and 400 ms. This is evident in the present data (particularly at TP10 in Figures 1c and 3c), and has also been observed very similarly in previous studies (see Figure 2c and 5c in Wiese, Tüttenberg, et al., 2019). Accordingly, these difference waves do not suggest a single effect that ramps up and then persists. Instead, they contain two peaks, with the second one being considerably larger than the first.

This raises the question *why* we reliably observe these two peaks in the difference waves. If visual face recognition is resolved in the N250 time window (e.g. Schweinberger & Neumann, 2016), why do we reliably observe the subsequent increase which peaks between 400 and 600ms? We have demonstrated in a previous study (Wiese, Ingram, et al., 2019) that the SFE is substantially reduced by manipulations that do not affect the N250, which may suggest that the two effects do not reflect the same, but at least partially different processes. We therefore suggest that the later part of the waveform (the SFE) reflects processing in addition to what we see in the earlier part (the N250 effect). At the same time, given the very similar occipito-temporal scalp distribution of the two effects, we do not assume that the N250 and the SFE are fully independent. Indeed, we have argued previously (Wiese, Tüttenberg, et al., 2019) that the SFE reflects feedback from later (affective, semantic, episodic) processing stages into earlier visual areas.

At the individual participant level, bootstrapping analysis indicated that the SFE was clearly more diagnostic of familiarity than the N250. Similarly, classifier performance increased over time and typically peaked after 400 ms. Accordingly, the diagnostic value of the signal gets stronger over time. Interestingly, however, reliable effects as observed using bootstrapping in individual participants were less likely to be observed in the concealed relative to the acknowledged familiarity condition in Experiment 3, with hit rates of .68 and .53, respectively, in the concealed condition. Moreover, a minority of participants in Experiment 2 ($P = .26$ for the N250 and $P = .16$ for the SFE) yielded false alarms, i.e., false familiarity effects when in fact the critical facial identity was unfamiliar. It thus appears that the CIT-like procedure examined in the present experiments is only moderately sensitive relative to previous findings. As detailed above, this interpretation is additionally supported by the results of the classifier analysis. Interestingly, using a butterfly detection task (Experiment 1), we observed a much higher hit rate of .96 in the bootstrapping analysis. This is consistent with our previous studies, reporting high hit rates with incidental recognition tasks. For example, Wiese et al. (2019) reported a hit rate of .85 while the false alarm rate was below .05 for the SFE. Even when participants were moderately distracted by task-relevant letter strings superimposed on the faces, we detected a hit rate of .83 for the N250 and of .78 for the SFE (Wiese, Ingram, et al., 2019). This latter finding appears particularly interesting, as the distracting task might hamper any deliberate attempt to suppress face recognition. In other words, when trying to find out whether a participant knows a critical facial identity or not, moderate distraction by an additional task might be more effective than inducing and then trying to detect active lying, as participants in the former scenario may not have the necessary resources to conceal familiarity. It is unclear, however, how likely false positive results are to occur in this experimental set-up. Moreover, the sensitivity of the procedure to the potential application of countermeasures (Rosenfeld et al., 2008; Rosenfeld et al., 2004) has not been tested as yet.

Having described the moderate sensitivity of the present approach, we will now discuss how it compares to alternative techniques. A P300-based CIT experiment reported by Meijer and colleagues (2007) observed concealed familiarity effects in 22 out of 24 participants (P = .92), which is comparable to the hit rate in Experiment 1 but substantially higher than in Experiment 2. However, the use of single images of facial identities in this study is potentially problematic, as a participant may be familiar with a specific picture without knowing the corresponding person. Moreover, the probability of false positive results, and hence the sensitivity of the procedure, remains unclear.

Using Fast Periodic Visual Stimulation (FPVS), Campbell and colleagues (Campbell, Louw, Michniak, & Tanaka, 2020) found reliable familiarity responses to personally familiar faces in eight out of twelve participants (P = .67), which is better than the concealed condition in Experiment 3 but lower than the typical hit rates for the SFE in butterfly detection tasks. Moreover, Yan and colleagues (Yan, Zimmermann, & Rossion, 2020) reported false positive familiarity responses in five out of 14 participants (P = .36) when faces were presented in upright orientation, and even in eight out of 14 participants (P = .57) with inverted stimulus presentation (i.e., in the condition in which face recognition should be less likely). A further recent study (Yan & Rossion, 2020) reports reliable familiarity effects in all tested participants, which is similar to the present Experiment 1, but does not report a false alarm rate. Given the relatively small sample size in all experiments, including those we present here, there remains substantial uncertainty about the precision of these estimates (Cumming, 2012). However, on the basis of the available data it seems fair to conclude that the sensitivity of FPVS to detect familiarity responses in individual participants is at best similar but probably lower relative to the SFE.

In conclusion, the present study used event-related potentials and a classifier based on supervised learning to investigate brain responses involving recognition of particular familiar faces. We were able to demonstrate that recognition is more consistently triggered when identification of a given person on the basis of a specific well-learnt image is not feasible. If this criterion is fulfilled, visual recognition, as reflected in the N250 familiarity effect, is independent of voluntary control and seems to occur largely automatically. Access to further identity-specific information, as represented by the SFE, similarly appears to be independent of the intention to recognise a person but depends on the availability of attentional resources. As the present findings also suggest that a CIT-like paradigm based on EEG/ERP correlates of face familiarity can detect concealed recognition with only moderate sensitivity, future studies should further explore ways to establish a viewer's familiarity with a particular person when moderately distracted from the face stimuli themselves.

References

Andrews, S., Burton, A. M., Schweinberger, S. R., & Wiese, H. (2017). Event-related potentials reveal the development of stable face representations from natural variability. *Quarterly Journal of Experimental Psychology, 70*(8), 1620-1632. doi:10.1080/17470218.2016.1195851

Bauer, R. M. (1984). Autonomic recognition of names and faces in prosopagnosia: a neuropsychological application of the Guilty Knowledge Test. *Neuropsychologia, 22*(4), 457-469. doi:10.1016/0028-3932(84)90040-x

Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological Studies of Face Perception in Humans. *J Cogn Neurosci, 8*(6), 551-565. doi:10.1162/jocn.1996.8.6.551

Bentin, S., & Deouell, L. Y. (2000). Structural encoding and identification in face processing: erp evidence for separate mechanisms. *Cogn Neuropsychol, 17*(1), 35-55. doi:10.1080/026432900380472

Bindemann, M., Burton, A. M., & Jenkins, R. (2005). Capacity limits for face processing. *Cognition, 98*(2), 177-197. doi:10.1016/j.cognition.2004.11.004

Bishop, C. M. (1995). Training with noise is equivalent to Tikhonov regularization. *Neural Computation, 7*(1), 108-116.

Bishop, C. M. (2006). *Pattern recognition and machine learning*. New York: Springer.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the Self-Assessment Manikin and the Semantic Differential. *J Behav Ther Exp Psychiatry, 25*(1), 49-59. doi:10.1016/0005-7916(94)90063-9

Bruce, V., Henderson, Z., Greenwood, K., Hancock, P. J. B., Burton, A. M., & Miller, P. (1999). Verification of face identities from images captured on video. *Journal of Experimental Psychology-Applied, 5*(4), 339-360. doi:10.1037/1076-898x.5.4.339

Bruce, V., Henderson, Z., Newman, C., & Burton, A. M. (2001). Matching identities of familiar and unfamiliar faces caught on CCTV images. *Journal of Experimental Psychology-Applied, 7*(3), 207-218. doi:10.1037/1076-898x.7.3.207

Bruce, V., & Young, A. (1986). Understanding face recognition. *Br J Psychol, 77 (Pt 3)*, 305-327. doi:10.1111/j.2044-8295.1986.tb02199.x

Bruce, V., & Young, A. W. (2012). *Face Perception*. Hove: Psychology Press.

Burton, A. M., Kramer, R. S., Ritchie, K. L., & Jenkins, R. (2016). Identity From Variation: Representations of Faces Derived From Multiple Instances. *Cogn Sci, 40*(1), 202-223. doi:10.1111/cogs.12231

Burton, A. M., Wilson, S., Cowan, M., & Bruce, V. (1999). Face recognition in poor-quality video: Evidence from security surveillance. *Psychological Science, 10*(3), 243-248. doi:10.1111/1467-9280.00144

Caharel, S., Courtay, N., Bernard, C., Lalonde, R., & Rebai, M. (2005). Familiarity and emotional expression influence an early stage of face processing: an electrophysiological study. *Brain Cogn, 59*(1), 96-100. doi:10.1016/j.bandc.2005.05.005

Caharel, S., Jacques, C., d'Arripe, O., Ramon, M., & Rossion, B. (2011). Early electrophysiological correlates of adaptation to personally familiar and unfamiliar faces across viewpoint changes. *Brain Res, 1387*, 85-98. doi:10.1016/j.brainres.2011.02.070

Campbell, A., Louw, R., Michniak, E., & Tanaka, J. W. (2020). Identity-specific neural responses to three categories of face familiarity (own, friend, stranger) using fast periodic visual stimulation. *Neuropsychologia, 141*, 107415. doi:10.1016/j.neuropsychologia.2020.107415

Cumming, G. (2012). *Understanding the New Statistics*. New York: Routledge.

Dalski, A., Kovacs, G., & Ambrus, G. G. Evidence for a general neural signature of face familiarity. *bioRxiv preprint*. doi:10.1101/2021.04.18.440317

Davies-Thompson, J., Gouws, A., & Andrews, T. J. (2009). An image-dependent representation of familiar and unfamiliar faces in the human ventral stream. *Neuropsychologia, 47*(6), 1627-1635. doi:10.1016/j.neuropsychologia.2009.01.017

Di Nocera, F., & Ferlazzo, F. (2000). Resampling approach to statistical inference: bootstrapping from event-related potentials data. *Behav Res Methods Instrum Comput, 32*(1), 111-119. doi:10.3758/bf03200793

Eimer, M. (2011). The Face-Sensitive N170 Component of the Event-Related Brain Potential. In A. J. Calder, G. Rhodes, M. H. Johnson, & J. V. Haxby (Eds.), *The Oxford Handbook of Face Perception*. Oxford: Oxford University Press.

Eimer, M., Gosling, A., & Duchaine, B. (2012). Electrophysiological markers of covert face recognition in developmental prosopagnosia. *Brain, 135*(Pt 2), 542-554. doi:10.1093/brain/awr347

Farwell, L. A., & Donchin, E. (1991). The truth will out: interrogative polygraphy ("lie detection") with event-related brain potentials. *Psychophysiology, 28*(5), 531-547. doi:10.1111/j.1469-8986.1991.tb01990.x

Gobbini, M. I., & Haxby, J. V. (2007). Neural systems for recognition of familiar faces. *Neuropsychologia, 45*(1), 32-41. doi:10.1016/j.neuropsychologia.2006.04.015

Gosling, A., & Eimer, M. (2011). An event-related brain potential study of explicit face recognition. *Neuropsychologia, 49*(9), 2736-2745. doi:10.1016/j.neuropsychologia.2011.05.025

Henderson, Z., Bruce, V., & Burton, A. M. (2001). Matching the faces of robbers captured on video. *Applied Cognitive Psychology, 15*(4), 445-464. doi:10.1002/acp.718

Jenkins, R., & Burton, A. M. (2011). Stable face representations. *Philos Trans R Soc Lond B Biol Sci, 366*(1571), 1671-1683. doi:10.1098/rstb.2010.0379

Jenkins, R., Burton, A. M., & Ellis, A. W. (2002). Long-term effects of covert face recognition. *Cognition, 86*(2), B43-52. doi:10.1016/s0010-0277(02)00172-5

Jenkins, R., White, D., Van Montfort, X., & Burton, A. M. (2011). Variability in photos of the same face. *Cognition, 121*(3), 313-323. doi:10.1016/j.cognition.2011.08.001

Johnston, P., Overell, A., Kaufman, J., Robinson, J., & Young, A. W. (2016). Expectations about person identity modulate the face-sensitive N170. *Cortex, 85*, 54-64. doi:10.1016/j.cortex.2016.10.002

Kaufmann, J. M., Schweinberger, S. R., & Burton, A. M. (2009). N250 ERP correlates of the acquisition of face representations across different images. *J Cogn Neurosci, 21*(4), 625-641. doi:10.1162/jocn.2009.21080

Kemp, R., Towell, N., & Pike, G. (1997). When seeing should not be believing: Photographs, credit cards and fraud. *Applied Cognitive Psychology, 11*(3), 211-222. Retrieved from <Go to ISI>://WOS:A1997XG45700002

Kramer, R. S. S., Young, A. W., & Burton, A. M. (2018). Understanding face familiarity. *Cognition, 172*, 46-58. doi:10.1016/j.cognition.2017.12.005

Kriegeskorte, N., Goebel, R., & Bandettini, P. (2006). Information-based functional brain mapping. *Proc Natl Acad Sci U S A, 103*(10), 3863-3868. doi:10.1073/pnas.0600244103

Lavie, N., Ro, T., & Russell, C. (2003). The role of perceptual load in processing distractor faces. *Psychol Sci, 14*(5), 510-515. doi:10.1111/1467-9280.03453

Logan, G. D. (1988). Automaticity, resources, and memory: theoretical controversies and practical implications. *Hum Factors, 30*(5), 583-598. doi:10.1177/001872088803000504

Meijer, E. H., Smulders, F. T., Merckelbach, H. L., & Wolf, A. G. (2007). The P300 is sensitive to concealed face recognition. *Int J Psychophysiol, 66*(3), 231-237. doi:10.1016/j.ijpsycho.2007.08.001

Meijer, E. H., Smulders, F. T., & Wolf, A. (2009). The contribution of mere recognition to the p300 effect in a concealed information test. *Appl Psychophysiol Biofeedback, 34*(3), 221-226. doi:10.1007/s10484-009-9099-9

Moors, A. (2016). Automaticity: Componential, Causal, and Mechanistic Explanations. *Annu Rev Psychol, 67*, 263-287. doi:10.1146/annurev-psych-122414-033550

Neumann, M. F., & Schweinberger, S. R. (2008). N250r and N400 ERP correlates of immediate famous face repetition are independent of perceptual load. *Brain Res, 1239*, 181-190. doi:10.1016/j.brainres.2008.08.039

Olivares, E. I., Iglesias, J., Saavedra, C., Trujillo-Barreto, N. J., & Valdes-Sosa, M. (2015). Brain Signals of Face Processing as Revealed by Event-Related Potentials. *Behavioural Neurology*, 1-16. doi:10.1155/2015/514361

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola Symposium* (pp. 55-85). Hillsdale: Lawrence Erlbaum Associates.

Ritchie, K. L., & Burton, A. M. (2017). Learning faces from variability. *Q J Exp Psychol (Hove), 70*(5), 897-905. doi:10.1080/17470218.2015.1136656

Rosenfeld, J. P., Cantwell, B., Nasman, V. T., Wojdac, V., Ivanov, S., & Mazzeri, L. (1988). A modified, event-related potential-based guilty knowledge test. *Int J Neurosci, 42*(1-2), 157-161. doi:10.3109/00207458808985770

Rosenfeld, J. P., Labkovsky, E., Winograd, M., Lui, M. A., Vandenboom, C., & Chedid, E. (2008). The Complex Trial Protocol (CTP): a new, countermeasure-resistant, accurate, P300-based method for detection of concealed information. *Psychophysiology, 45*(6), 906-919. doi:10.1111/j.1469-8986.2008.00708.x

Rosenfeld, J. P., Soskins, M., Bosh, G., & Ryan, A. (2004). Simple, effective countermeasures to P300-based tests of detection of concealed information. *Psychophysiology, 41*(2), 205-219. doi:10.1111/j.1469-8986.2004.00158.x

Rossion, B., & Jacques, C. (2008). Does physical interstimulus variance account for early electrophysiological face sensitive responses in the human brain? Ten lessons on the N170. *Neuroimage, 39*(4), 1959-1979. doi:10.1016/j.neuroimage.2007.10.011

Saavedra, C., Iglesias, J., & Olivares, E. I. (2010). Event-Related Potentials Elicited by the Explicit and Implicit Processing of Familiarity in Faces. *Clinical Eeg and Neuroscience, 41*(1), 24-31. doi:10.1177/155005941004100107

Schweinberger, S. R., & Neumann, M. F. (2016). Repetition effects in human ERPs to faces. *Cortex, 80*, 141-153. doi:10.1016/j.cortex.2015.11.001

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information-processing II: Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84*(2), 127-190. doi:10.1037/0033-295x.84.2.127

Tanaka, J. W., Curran, T., Porterfield, A. L., & Collins, D. (2006). Activation of preexisting and acquired face representations: the N250 event-related potential as an index of face familiarity. *J Cogn Neurosci, 18*(9), 1488-1497. doi:10.1162/jocn.2006.18.9.1488

Tranel, D., Damasio, H., & Damasio, A. R. (1995). Double Dissociation between Overt and Covert Face Recognition. *J Cogn Neurosci, 7*(4), 425-432. doi:10.1162/jocn.1995.7.4.425

Treder, M. S. (2020). MVPA-Light: A Classification and Regression Toolbox for Multi-Dimensional Data. *Front Neurosci, 14*, 289. doi:10.3389/fnins.2020.00289

Weibert, K., Harris, R. J., Mitchell, A., Byrne, H., Young, A. W., & Andrews, T. J. (2016). An image-invariant neural response to familiar faces in the human medial temporal lobe. *Cortex, 84*, 34-42. doi:10.1016/j.cortex.2016.08.014

White, D., Kemp, R. I., Jenkins, R., Matheson, M., & Burton, A. M. (2014). Passport officers' errors in face matching. *PLoS One, 9*(8), e103510. doi:10.1371/journal.pone.0103510

Wiese, H., Ingram, B. T., Elley, M. L., Tüttenberg, S. C., Burton, A. M., & Young, A. W. (2019). Later but not early stages of familiar face recognition depend strongly on attentional resources: Evidence from event-related brain potentials. *Cortex, 120*, 147-158. doi:10.1016/j.cortex.2019.06.004

Wiese, H., Tüttenberg, S. C., Ingram, B. T., Chan, C. Y. X., Gurbuz, Z., Burton, A. M., & Young, A. W. (2019). A Robust Neural Index of High Face Familiarity. *Psychological Science, 30*(2), 261-272. doi:10.1177/0956797618813572

Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010). Controlling low-level image properties: the SHINE toolbox. *Behav Res Methods, 42*(3), 671-684. doi:10.3758/BRM.42.3.671

Yan, X., & Rossion, B. (2020). A robust neural familiar face recognition response in a dynamic (periodic) stream of unfamiliar faces. *Cortex, 132*, 281-295. doi:10.1016/j.cortex.2020.08.016

Yan, X., Young, A. W., & Andrews, T. J. (2017). The automaticity of face perception is influenced by familiarity. *Atten Percept Psychophys, 79*(7), 2202-2211. doi:10.3758/s13414-017-1362-1

Yan, X., Zimmermann, F. G., & Rossion, B. (2020). An implicit neural familiar face identity recognition response across widely variable natural views in the human brain. *Cogn Neurosci, 11*(3), 143-156. doi:10.1080/17588928.2020.1712344

Young, A. W., & Burton, A. M. (2018). Are We Face Experts? *Trends Cogn Sci, 22*(2), 100-110. doi:10.1016/j.tics.2017.11.007

Young, A. W., Hellawell, D., & De Haan, E. H. (1988). Cross-domain semantic priming in normal subjects and a prosopagnosic patient. *Q J Exp Psychol A, 40*(3), 561-580. doi:10.1080/02724988843000087