Brain Informatics

RESEARCH

Open Access

Breakthrough percepts of familiar faces

Omid Hajilou[1] and Howard Bowman[2,3*]

Abstract
In Rapid Serial Visual Presentation (RSVP), the vast majority of stimuli are not consciously perceived, but the salient ones breakthrough into awareness and can be reported. In addition, these breakthrough events are observable with EEG, since they generate a P3 or other distinguishing components. The Fringe-P3 method is based upon these characteristics. Concealed knowledge studies have successfully employed this Fringe-P3 method using own-name and own email-address, with the method being shown to be less vulnerable to counter-measures than other approaches. It has also been shown that famous faces presented in RSVP differentially break into awareness and generate a distinct evoked response component. In this paper, we further enhance the applicability of the Fringe-P3 method by demonstrating the effectiveness of the method on personally-familiar faces. While salient, such stimuli do not have the exquisite salience of famous faces, being a better match to the level of salience that might be found in forensic applications. Our findings suggest that the Fringe-P3 method could be used to detect intrinsic salience of familiar faces, even when there was no task associated with these faces. We investigated the sensitivity of the ERP-based RSVP paradigm to infer recognition of familiar faces, and performed statistical inference in the Time and Frequency domains, to differentiate between known and unknown faces, at group and participant levels.

Keywords Lie detection, Event related potentials, Concealed knowledge test, Fringe-P3 method, EEG, Time–frequency analysis, Familiar faces, Rapid serial visual presentation

1 Introduction

The classic oddball P3 experiments have been successfully used in an investigative context as Concealed Information Tests (CIT) [29, 33, 41]. Additionally, CITs have successfully used face stimuli in these P3-oddball paradigms [34, 37]. However, these P3-oddball paradigms have exhibited a vulnerability to countermeasures [41]. Recent CIT work with Rapid Serial Visual Presentation (RSVP) [12, 13] has been argued to be less vulnerable to countermeasures. In particular, salient stimuli that are presented in RSVP streams (e.g. a familiar name) will breakthrough into awareness, whilst non-salient stimuli (e.g. an unknown name) will not enter awareness making it more difficult for participants to use countermeasures. For example, where stimuli are presented at a slow rate, suspects could confound the tests by using countermeasures [35, 38]. A key reason for the possibility of countermeasures is the availability of sufficient time between each stimulus, to consciously determine that there is a repeating non-critical irrelevant item.[1] However, [12] observed that if the stimuli were presented at a rapid rate in RSVP, the irrelevant stimulus was not sufficiently perceived to be noticed as repeating, preventing it being

*Correspondence:
Howard Bowman
H.Bowman@bham.ac.uk
[1]School of Computing, University of Kent, Canterbury, UK
[2]School of Psychology and School of Computer Science, University of Birmingham, Birmingham B15 2TT, UK
[3]Functional Imaging Lab, University College London, London, UK

[1] If the Irrelevant can be detected, the electrical response to it can be increased, potentially making its response statistically indistinguishable from that to the concealed knowledge (the Probe), confounding the deception detection.

used in a countermeasure. This difficulty for participants to detect non-salient repeating stimuli has been confirmed in a range of empirical work [3, 8, 9].

These RSVP findings were implemented in the Fringe-P3 method, a brain computer interface, which emerged from theoretical work on the attentional blink [10, 14, 18]. The method involves presenting an RSVP stream and determining the stimuli in the stream that the participant finds salient by detecting a "breakthrough" event, i.e. the participant's perceptual system detecting a salient stimulus and accordingly its representation being consciously perceived. This perceptual event manifests as a P3 or other distinguishing EEG component. The method has been used as a P3-speller, a classic brain computer interface [17] and as a CIT.

Having provided evidence for the effectiveness of the Fringe-P3 method based upon own-name [12, 13], own email-address [25], famous name [2], future experiments were required to explore the suitability of image-based stimuli, in ERP-based RSVP paradigms, for example, do familiar faces differentially break into conscious awareness, on an individual basis, and can we detect the breakthrough events with EEG? To answer this question, we began by studying the effects of presenting famous faces (i.e. Probe stimuli and unknown faces (i.e. Irrelevant stimuli, using the RSVP technique.

Accordingly, we are able to successfully establish that famous/celebrity faces breakthrough into conscious awareness, using an RSVP subliminal search paradigm, and that our statistical tests can differentiate between the Probe (famous) and Irrelevant (unknown) faces, at group and participant levels [1]. The objective of the current paper is to demonstrate that we can substitute the highly evocative faces of famous celebrities with familiar faces that are personally known to the participants. While salient, such stimuli do not have the exquisite salience of famous faces, and are thus a better match to the level of salience that might be found in forensic applications. Additionally, we assessed whether the Fringe-P3 method could be used to detect intrinsic salience of familiar faces, even when there was no task associated with these faces. (Although, to confirm the applicability of this work for a CIT, more research verifying this familiar-faces finding in more ecologically valid deception detection contexts needs to be undertaken, e.g. with participants attempting to apply countermeasures.)

ERP studies into face perception and recognition have reliably reported neural activity, with specific components that indicate sub-conscious and conscious processing of faces. Numerous studies have reported a face-specific N170 component—a negative deflection, elicited in the ERP within 140 and 200 ms post-stimulus.

Whilst the N170 component does not appear to be affected by the difference between famous and unfamiliar faces [6], an enhanced negativity called the N400 component (also referred to as N400face, or N400f) appears to be associated with the subsequent activation of the episodic memory of the face [6, 23, 46]. Amongst others, [23] observed a subsequent/late positivity for famous faces, referred to as the P600 component (or P600f), which has also been compared with the P3 component [45]. This P3/P600f neural response may reflect explicit recognition of a particular individual.

In Alsufyani et al. [1], we used the N400f and the P600f for our time domain analyses of the brain response to famous faces presented in RSVP both at the group and per-individual level. However, our most effective per-individual statistical test in Alsufyani et al. [1] was in the frequency domain, where we performed a permutation test over power and inter-trial coherence features. These time and frequency domain analyses will be reused in this paper exactly as used in Alsufyani et al. [1] to differentiate between the Probe (familiar: lecturer at participant's institution, University of Kent) and Irrelevant (unknown: lecturer at different institution, Christchurch University). This gives us strong prior precedents for the analyses we perform here, providing robust protection against type 1 errors.

The aim of this work is, then, to provide a proof of concept, which could be further refined and developed into a scientifically robust framework, in the pursuit of a means by which a suspect's familiarity with compatriots can be demonstrated using EEG.

2 Methods

2.1 Participants

Fourteen participants were tested, who were all students at the University of Kent, free from neurological disorders, and with normal, or corrected-to-normal vision. None were excluded. Out of 14 participants, 12 were male (86%) and 2 female (14%). The ages of the participants ranged from 22 to 37 ($M = 27.5$ years, $SD = 3.94$); 12 of them were right-handed (86%), and 2 were left-handed (14%). Because this experiment's Probe stimuli consisted of University of Kent's School of Computing lecturer faces, we needed participants who had a close working relationship with these lecturers. Therefore, we asked the lecturers to covertly nominate PhD students, so that we could be assured of a long-term relationship/familiarity between participants and their lecturers' faces. Also, participants were chosen on the basis of never having been included in a similar EEG/RSVP experiment, and at the end of each experiment, participants were instructed to avoid discussing the experiment with their colleagues, in order to avoid any priming of future participants. The duration of each experiment was (approx.) 1 h and 45 min, and each participant was paid £10 (ten pounds) for their time. The Sciences Research Ethics Advisory Group, at the University of Kent, approved the study.

2.2 Stimuli

The instructions, stimuli and questions were presented on a 20-inch LCD monitor, with a refresh rate of 60 Hz, and a resolution of 1600 × 1200 pixels. The screen was placed at a comfortable position for each participant, at a distance of approximately 60 to 80 cm. All stimuli were scaled to 280 × 320 pixels, and presented in the centre of the screen, using the Rapid Serial Visual Presentation (RSVP) method.

The stimuli were split into two groups: a) *Distractor* images (i.e. unknown faces) and b) *Critical* images, as described below:

a. *Distractor* images (i.e. the first group) were photographs of unfamiliar faces, which were obtained from an open-source, online database of faces [39] from the University of Texas at Dallas – all of these faces were frontal views.
b. *Critical* images (i.e. the second group) consisted of the following three categories:
    1. *Target* face, which was a single image, chosen by us from the Distractors database. The Target was task-relevant (i.e. participants were instructed to look out for, and respond to this image).
    2. *Irrelevant* (a.k.a. *unknown*) faces consisted of unknown faces of Lecturers from another University.
    3. *Probe* (a.k.a. *familiar*) faces consisted of several images of familiar faces, in the form of our University's lecturer faces.

Care was taken to ensure that all the images used in the experiment would conform to our compatibility criteria. Accordingly, images with incongruous features (e.g. angry or smiley faces), which could breakthrough into conscious awareness due to their dissonant features, were avoided. As a result, the large database of Distractor faces (with over 1000 images) was reduced to 573 'neutral' images, without significant facial expressions or features. The remaining 573 Distractor images were available as fillers for RSVP streams, albeit, one was selected as the Target, for the experiment. To be clear, this exclusion of stimuli with significant facial expressions, cannot impact the inferences made in this paper, since this was only performed for face stimuli that would become Distractors in the experiment. In particular, the resulting set of faces was sampled from randomly and statistically identically for streams containing Probes and those containing Irrelevants. Accordingly, this selection cannot confound the logic of our experiment, since if Distractors are somehow more or less attention grabbing, they will be equally so in the context in which Probes arise as in the context in which Irrelevants arise.

The Probes (i.e. familiar Lecturer faces from the University of Kent's School of Computing), as well as, the Irrelevants (i.e. unknown Lecturer faces from Christ Church University) were taken using the same SLR camera (Canon PowerShot G11). All lecturers consented to their image being used in our EEG experiments; these photos were taken from the same position/distance, under similar lighting conditions, and with neutral poses.

All images of lecturers were manually edited, in order to remove any non-conforming distinguishing features, and to obtain similar brightness and contrast. First, each image was centred by aligning the eye-line to the same horizontal position, and then resized, to occupy the same space/size as the Distractor images. Next, the background of each lecturer face was removed (i.e. borders were carefully highlighted/selected, using the photo-editing tool GIMP, before being cropped out), and then the background colour was changed to light-grey (Hex colour: #e7e7e7). Next, the contrast of the images were reduced, wherever necessary, and all Probe images were resized to 280 × 320 pixels and converted to monochrome. (Appendix 1 contains a systematic comparison of the low-level visual features of Probe and Irrelevant images, which suggests that there are not differences between the two classes of image that could explain the effects we observe.)

After the above exercise to select-and-edit our Probe images, we decided to further reduce the Distractor images to 524 faces, in order to approximately match the age range of the Probe faces (i.e. by excluding Distractor images of very young and very old individuals). As there were 524 possible Distractors, which could be used as fillers for RSVP streams, the probability that one would be randomly selected for each stream was 0.032, and equal for all of them.

2.3 Probe/irrelevant comparison

As explained above, Probe and Irrelevant critical stimuli were lecturer faces that were photographed by us. Having photographed a large portion of the lecturers in the University of Kent's School of Computing (23 images in total), we were able to assign three Lecturers (as Probes) to each participant, knowing that they were familiar with one another (as confirmed by the Lecturers and/or the participant's colleagues). Additionally, each participant was randomly assigned three unknown lecturers (as Irrelevants) from a different University (i.e. Canterbury Christ Church University).

2.4 Design

All stimuli were presented using the *Psychophysics* toolbox version 3, running under MATLAB 2012a. All RSVP stream items were presented in the same location (centre of the screen), at an SOA of 133 ms (see Fig. 1), and without an Inter-stimulus-Interval (ISI). Each RSVP stream contained 18 faces, 17 of which were Distractors (i.e. unknown fillers), and only one was a Critical item (i.e. either a Probe, Irrelevant or Target). To ensure that neither the Probes nor the Irrelevants were task-relevant, we instructed all participants to look for a single Target (i.e. an unknown face that participants were trained to detect), throughout the experiment. This task-relevant Target was repeated as many times as each of the other two Critical stimuli (i.e. the Target, Probe and Irrelevant conditions were repeated equal number of times) and (in a statistical sense) in the same position in streams.

We implemented a change to the design of the famous faces experiment [1]: to improve the signal-to-noise ratio of unique Probe/Irrelevant Critical stimuli, we reduced the number of Probes/Irrelevants from five (celebrities) to three (lecturers), whilst maintaining the total number of trials for each participant (i.e. 225 trials in total, so that the number of trials per Critical condition would also remain the same between the two experiments). Therefore, in this experiment, each Probe was repeated 25 times, resulting in 75 Probe-trials (i.e. 25 times for each of the 3 Probes), and each Irrelevant was also repeated 25 times, resulting in 75 Irrelevant-trials (i.e. 25 times for each of the 3 Irrelevants). The single Target was also repeated 75 times, to equal the number of repetitions of the other two Critical Stimuli. The resultant 225 RSVP trials were divided into 3 blocks, each block comprising 75 trials (i.e. 25 Probe trials, 25 Irrelevant trials, and 25 Target trials), and the order of the three Critical stimuli were randomised within the blocks. However, each block's Probe and Irrelevant Critical stimuli were paired, so that the same known lecturer (Probe) face and unknown lecturer (Irrelevant) face were presented within the same block—this will enable us to make direct comparisons between these paired-conditions.

Participants were told to keep their eyes fixed at the centre of the screen during the presentation of the RSVP stream (lasting approx. 2.5 s), and to avoid movement or blinking. Also, they were informed that the Target image will appear pseudo-randomly, so they should not expect it in every trial, however, participants were naïve to the presence of familiar lecturer faces (i.e. Probes).

In this experiment, out of a total of 75 trials for each Critical condition, the number of trials that remained after artefact rejection, per condition, ranged between 59 and 75, and none of the participants were excluded from the analysis due to removal of artefact trials (e.g. excessive eye blinks):

$$Target\ (M = 70.5,\ SD = 4.26);$$

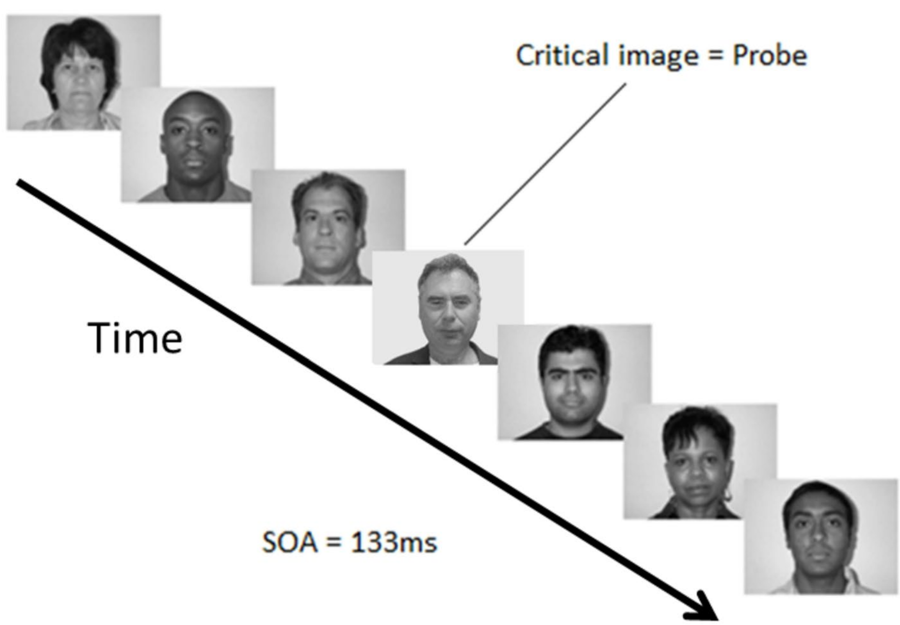

Fig. 1 RSVP stream, showing 7 of the 18 faces that could be presented in a trial, where each trial consists of 17 Distractors (i.e. unfamiliar faces), and one Critical image (i.e. a Probe, Irrelevant or Target). In the above example from the experiment, a Lecturer (Howard Bowman) is the Probe that is presented as the Critical image

$$Probe\ (M = 71.93,\ SD = 3.27);$$

$$Irrelevant\ (M = 72.64,\ SD = 2.17).$$

The position of the single Critical item within each RSVP trial, which contained 17 Distractors as fillers, was selected pseudo-randomly by the application, so that it had equal probability of appearing anywhere in the 5th to 9th position of the stream of images. This choice to restrict critical stimuli (Targets, Probes, Irrelevants) to positions 5–9 in RSVP streams, which is consistent with our previous experiments (e.g. [1]), is not arbitrary. In fact, this restriction is required to keep the early deflections associated with a critical stimulus free of contamination from the transients initiated by the start of the stream and later deflections free of contamination from offset transients and anticipation of the end of stream item-and-question. Restricting presentation of critical stimuli to particular serial positions could induce temporal expectations, but since we also do this for Irrelevants, it cannot break the logic of our experiment.

In addition to the 18 images, each RSVP stream contained a starting and finishing item, which would improve the participants' focus, from the beginning to the end of each stream (i.e. to know when the stream is about to start, and when it will end). A starting item "+ + + + + + +" was presented for 800 ms, to position the participant's focus on the presentation area of the screen. After the last image of the RSVP stream (i.e. the 18th face), a random finishing item (a.k.a. attention-checker image), which could either be "- - - - - - -" or "= = = = = = =", was presented for 133 ms; this end-of-stream image required the participant to remain attentive until the end of the stream. Using a standalone keypad, which was positioned under the participant's preferred left or right hand, the participant was asked to report the attention-checker item, using '1' and '2' keys (1 signifying "- - - - - - -" and 2 signifying "= = = = = = =").

As soon as the participant responded to the attention-checker question, a task-relevant question was shown, in order to determine the detection (or not) of the Target image, using the question: "Did you see the *Target* image within the stream?". In response, the participant could select the keys '4' (for "Yes") or '5' (for "No"). The above two questions were repeated at the end of each RSVP stream. Recognition-questions were also included at the end of the experiment. These were included at experiment end in order to mitigate the participants' inference that the Probe (which would have been in the recognition test) could also be task-relevant (i.e. in addition to the Target), which could have occurred if recognition was tested during the course of the experiment.

Participants were given one (or more) training session(s). Each training session consisted of 20 RSVP trials, to ensure that the participant was comfortable observing the rapid presentation of images, and to make sure that they could identify the Target. During the training session(s), the RSVP streams did not contain any images that were assigned to the Probe or Irrelevant category of critical stimuli in the experiment proper.

2.5 Analysis

2.5.1 Data acquisition

The experiment was recorded using a BioSemi ActiveTwo system (*BioSemi, Amsterdam, The Netherlands; see* www.biosemi.com). The Electroencephalographic (EEG) data was filtered at recording, with a low-pass of 100 Hz, and digitised at 2,048 Hz, for offline analysis, and the impedance was kept below 10 kΩ. In accordance with the standard 10–20 system [28], the following 8 scalp electrodes were used in the first two experiments: Fz, Cz, Pz, P3, P4, Oz, A1 and A2. During recording, data was referenced to a ground formed from a Common Mode Sense (CMS) active electrode, and Driven Right Leg (DRL) passive electrode. These two electrodes form a feedback loop, which drives the average potential of the participant, as close as possible, to the Analogue to Digital Converter (ADC). Also, Electrooculograms (EOG) generated from eye blinks and eye movements, were recorded from the participant's left and right eyes, using two bipolar Horizontal EOG (HEOG) and Vertical EOG (VEOG) electrodes.

2.5.2 EEG data

For all three experiments, the recorded data, which was analysed using EEGLAB (version 12.02.4b), under MATLAB 2012a [21], was resampled at 512 Hz. Before analysis, we filtered the data with low-pass, and high-pass, filters of 45 Hz to 0.5 Hz, respectively. Furthermore, in order to remove the steady state Visually Evoked Potential (ssVEP) oscillations [48], notch filters were applied between 7 and 9 Hz. Then, we re-referenced the data off-line to the average of the combined mastoids (i.e. A1 and A2 electrodes), and generated ERPs by separately averaging all trials for each Critical condition (i.e. Target, Probe, and Irrelevant). Each EEG trial was generated by epoching the data, using a −200 ms to 1200 ms stimulus-locked window (i.e. −200 ms before the appearance of the Critical stimulus, and 1200 ms after the occurrence of the Critical stimulus), and all ERPs were time-locked to the onset of the Critical item (i.e. time-point zero marks the appearance of the Critical face). Although baseline correction could be applied at trial level (i.e. mean of −100 ms to 0 ms window subtracted from each trail), the new standard for applying the detrending technique (see Application of Detrending) required us to baseline correct each trial after the adjustment for any errant drift.

Eye blinks and muscle movements were detected, by marking any activity below − 100 μV or above + 100 μV in the EOG channels (reflecting eye blinks and horizontal/vertical movements). Furthermore, we rejected any trials containing electrical activity below − 50 μV or above + 50 μV, in a time window from − 200 ms to 1200 ms (reflecting other physiological and environmental artefacts), with respect to the Critical stimulus Onset. Finally, we performed manual inspection of the resulting ERP data, to verify that the rejected trials were accurately detected.

2.5.3 Time domain (ERP) analysis

Each experiment's trials can be split into three categories: those with the *Target* critical stimulus, ones with *Probe* critical stimulus and the ones with *Irrelevant* critical stimulus (which was paired with the Probe). Our analyses were performed at the ERP-level, in order to compare the responses to familiar (Probe) and unfamiliar (Irrelevant) faces. As previously discussed, we focused on two face related ERP components: 1) an enhanced negativity called N400—also referred to as N400f (Curan & [19]), and 2) a late positivity called P3—also referred to as P600f [47].

The time window associated with the above two ERP components should be identified based on a contrast independent of (strictly parametrically contrast orthogonal to) the contrast of interest [11, 15]. Thus, for pre-participant analysis, we selected the time window using the participant's aggregated ERP, generated from all trials (hereafter called aERPt, within the combined Probe-and-Irrelevant conditions; this is also called the aggregated ERP of trials [11, 15]. Similarly, for group-level analysis, we selected the time window using an Aggregated Grand Average from all Trials (AGAT), belonging to all participants' combined Probe-and-Irrelevant conditions [11, 15]. The following gives more details of the aERPt and AGAT.

2.5.4 Per-participant (aERPt) window placement and inference

For each participant, their aggregated ERP of all trials from both conditions, Probe and Irrelevant, was generated. This aERPt was then used to identify the time window of the two components of interest (i.e. N400f and P600f). For the N400f component, the entire ERP window (i.e. from 0 to 1200 ms) was searched to find the *minimal* 100 ms interval average. Similarly, for the P600f component, the entire ERP window (i.e. from 0 to 1200 ms) was searched to find the *maximal* 100 ms interval average. (We searched in the entire window, in order to give as much flexibility as possible for the latency of participant's brain responses to vary and also, to bring as few (potentially biasing) constraints as possible to our analysis.) The start and the end of this minimal/maximal 100 ms Region of Interest (ROI) defined the face related N400f and P600f time features/components.

After identifying the time windows for each component (e.g. the ROI for N400f, and the ROI for P600f), the mean amplitude was calculated separately to each condition within these time windows—in other words, for each component, one mean amplitude value for the Probe, and another for the Irrelevant, was calculated using the same time window that was independently found when Probe and Irrelevant trials were combined. The 'True Observed' difference of each component (i.e. N400f and P600f) in their respective ROI, is the difference between this measure for each condition:

$$
\begin{aligned}
&True\ Observed\ difference \\
&\quad = mean\ of\ Probe\,(minus)\,mean\ of\ Irrelevant
\end{aligned}
$$

Having found the True Observed difference for N400f and P600f, we performed statistical inference, to determine whether the evoked response by the Probe (the familiar face) was significantly different from that evoked by the Irrelevant (the unknown face). Per-individual analysis was based on analysing each experimental participant separately to determine whether there was a significant difference for that participant alone. In this analysis, a randomisation (i.e. Monte Carlo Permutation) test was used to determine a p-value for each participant (see section "Randomisation Test", below). A null hypothesis distribution for each participant was generated in order to calculate the individual's p-value; the p-value estimates the probability that the observed pattern could have arisen if the null hypothesis was true, i.e. the participant was not familiar with Probe face.

2.5.5 Group-level (AGAT) window placement and inference

Time window placement for the group starts with the collation of all the trials for all participants, in both conditions (Probe and Irrelevant). The resultant Aggregated Grand Average of Trials (AGAT) was then used to identify the time window in the entire ERP segment of the two components of interest (i.e. N400f and P600f). For the N400f component, the *minimal* 100 ms interval average was found, and for the P600f component, the *maximal* 100 ms interval average was found. The start and the end of this minimal/maximal 100 ms Region of Interest (ROI) defined the group-level face related N400f and P600f features/components. Finally, the mean amplitude measure was calculated separately for each condition, within the identified time window, and the 'True Observed' difference of each component (i.e. N400f and P600f) was obtained by finding the difference between each condition (similarly to in section "Per-participant (aERPt) window placement and inference").

Having found the True Observed difference for N400f and P600f, we performed statistical inference, to determine whether the evoked response for the Probe (the familiar face) was significantly different from that evoked for the Irrelevant (the unknown face). A paired *t*-test of the mean amplitudes of Probe N400f/P600f and Irrelevant N400f/P600f was then run across participants.

2.5.6 Randomisation test

At per-individual-level, to determine whether the difference between two conditions (Probe and Irrelevant) is significant, we applied a randomisation (i.e. Monte Carlo Permutation) test. This was done separately for N400f and P600f components, for each participant. Before applying the test at the per-participant-level, the smallest number of trials between the Probe and Irrelevant conditions was determined and denoted by '*m*' (they could contain different numbers of trials due to artefact rejection). Thus, only '*m*' trials were selected (at random without replacement) from the Probe condition, and '*m*' trials from the Irrelevant condition. Notably, if a direct comparison was to be made between individual blocks (as relating to a single familiar face), we made sure that the pairing of the Probe and Irrelevant conditions were maintained. Next, we calculated the difference between the mean amplitude values of Probe and Irrelevant ERPs (*Probe [minus] Irrelevant*), in order to obtain a mean amplitude difference measure. This mean amplitude difference became the True Observed Value.

The randomisation test was applied by populating a matrix of size ($2.m \times$ *number of time points*) with $2.m$ selected trials; row position was randomised in the matrix. Under the null hypothesis, the Irrelevant and Probe trials are samples from the same distribution (i.e. the null distribution), and would thus be exchangeable. This justifies the randomisation of position in the matrix. Next, a pair of datasets were generated: the first, the surrogate Probe, was generated from the first half of the matrix rows, and the second, the surrogate Irrelevant, was generated from the remaining half. The desired analysis (i.e. mean amplitude measure) was then applied to each of the two randomised data sets, and the mean amplitude value was calculated (referred to as *Surrogate Values*), in the same way that the True Observed Value was calculated.

The above randomisation procedure was repeated 1,000 times. In each iteration, a new mean amplitude difference was obtained, resulting in 1,000 *Surrogate Values*. The *p*-value was then calculated as the proportion of randomised results that were greater than the true observed value. Finally, if this *p*-value is smaller than a critical alpha-level (0.05), then the data in the two experimental conditions are significantly different, and the null Hypothesis (H0) can be rejected. Note that in each resampling, the randomised mean amplitude difference (i.e. surrogate value) was measured for both the N400f and the P600f components (i.e. these values were calculated from the same random sample, rather than being calculated from two separate randomisations).

2.5.7 Frequency domain analysis

Previous studies into familiar faces [6] have reported ongoing oscillations in ERPs, from about 100 ms to 500 ms post stimulus onset, over parietal and occipital sites. In our famous faces experiment [1], we also observed multi-cycle oscillations in the Probe, which are not present in the Irrelevant, and, also, not in the Target. Because classic ERP analysis methods, like peak-to-peak or base-to-peak, would not fully reflect or measure the Probe's multi-cycle oscillations, time frequency analyses (ERSP and ITC) were used, as outlined below.

To analyse EEG data in the time–frequency domain, the following two transforms were used: Event Related Spectral Perturbations (ERSP) and Inter-Trial Coherence (ITC). These were calculated, using a fast Fourier transform, with a baseline correction at – 100 ms to 0 ms. ERSP calculates the average changes, relative to baseline, in power at each frequency at each time point, across all trials [21]. ITC measures phase consistency between trials, determining the extent to which trials are phase-locked, at each time point and frequency [36].

2.5.8 Time frequency window placement

Time frequency analyses were measured, over two time windows, using orthogonal contrast time window placement [11, 15, 31, 32]. In a similar way to our ERP analysis of the time domain, the time window for the Region of Interest that we used to measure ERSP and ITC, was identified based on aggregated power and coherence.

For group-level Time Frequency analysis, the placement of the critical time window (i.e. the highest 100 ms interval in the broader time window of 0 ms to 1200 ms) for measuring ERSP/ITC was calculated using the average of power/coherence of all single trials of all participants (i.e. the aggregated grand average of all trials, across all participants) from both Probe and Irrelevant conditions. For participant-level Time Frequency analysis, the placement of the critical time window (i.e. the highest 100 ms interval in the broader time window of 0 ms to 1200 ms) for measuring ERSP/ITC was calculated using each individual participant's average of power/coherence (i.e. the aggregated ERP of all trials, for a single participant) from both Probe and Irrelevant conditions. Thus, both methods for time window placement were calculated independently of the contrast that is statistically tested.

Next, these orthogonally derived time windows could be employed to measure ERSP and ITC separately, in the Probe and Irrelevant conditions. In keeping with previous studies [21], the EEGLAB time–frequency function *newtimef* was used to calculate the ERSP and ITC for each condition. Each condition (i.e. Probe and Irrelevant) would supply this function with a matrix that contains its respective time-points and trials. The *newtimef* function would process these two input matrices and calculate two output matrices, which represent the difference in the power and coherence; the first output matrix comprised the difference in power (i.e. ERSP) between Probe and Irrelevant conditions, and the second comprised the difference in coherence (i.e. ITC) between Probe and Irrelevant conditions.

2.5.9 Time frequency statistical inference

We summed across all values that were greater than zero, in the (orthogonally-derived) time window for the frequency range of interest, to give us a measure for each of the ERSP and ITC transforms for each condition. For this summation process we derived two difference (Probe minus Irrelevant) measures—one for power (i.e. ERSP) and another for coherence (i.e. ITC)—which would become the True Observed Values. Just as we had done in the time domain, we used a randomisation (Monte Carlo permutation) procedure to generate two Null hypothesis distributions: one for power and one for ITC (as calculated by the summation process, outlined above, across the orthogonally derived time windows). For each participant, we calculated p-values for both power and ITC transforms.

2.5.10 Detrending

Drift in EEG data can be a significant problem for the selection of an analysis window based upon identifying an interval with an extreme mean amplitude. For example, if the signal is drifting upwards, it is likely that the most positive-going window is at the end of the time series, and this selection would be made because of the drift rather than the amplitude of any particular component in the time series. Thus, in order to successfully use the aERPt or AGAT approaches, one needs a means of mitigating drift. The standard method for dealing with drift in EEG data would be to employ increasing high-pass filtering. However, filtering strategies may introduce temporal distortions in the signal, especially for low frequency P3 components that we are studying, where the amplitude of the P3 starts to reduce as the frequency cut-off increases. Therefore, we sought a method for mitigating drift, without introducing temporal artefacts. This was achieved through a detrending technique, whereby a linear trend line was fitted to the combined probe/irrelevant average and then this line was subtracted from every trial of both conditions.

Finally, baseline correction was performed after detrending, otherwise, we could be artificially lowering the late-components and increasing the early-ones (i.e. tilting the data, end-down and start-up).

3 Results

3.1 Target question

As previously explained, at the end of each RSVP stream, the participant was required to answer two questions. The first question related to the finishing-item and the second related to Target-recognition. Out of 75 times that each participant was randomly presented with the Target face, the average Hit rate for the group was 72.6% ($M = 54.43$, $SD = 17.87$), and out of the remaining 150 other trials in which the Target was not presented, the False-Positive rate was 3.3% ($M = 4.93$, $SD = 4.12$). The resulting discrimination measure ($d' = 2.494$) suggested that participants performed the Target task well, and no participants were excluded due to low discrimination.

3.2 Probe/irrelevant questions

As previously discussed, at experiment end, the participant was given a recognition memory test, to determine if the three Probes and/or the three Irrelevants were perceived/recognised. This memory test consisted of 12 questions, appearing randomly, where each question accompanied an image that may or may not have been included in the experiment's three blocks. Six questions related to the presence of the paired Probe and Irrelevant Lecturer faces that were included in the three blocks (i.e. one pair of Probe/Irrelevant, per block), and the other six questions related to random Probe and Irrelevant (Lecturer) faces that were not included in the experiment.

As before, the response to each of these 12 recognition/memory tests were handled in two parts: firstly, what is the participant's confidence rating of how often each face was presented (i.e. the Probe/Irrelevant that were present, and the Probe/Irrelevant that were absent), and secondly, a confidence rating of how well the participant knew each of the 12 faces, prior to the experiment. The responses to both of these confidence ratings used a scale of 1 to 5, where 1 is "Never", 2 is "Once or twice", 3 is "Few times", 4 is "Many times" and 5 is "A lot". Note, for the purposes of statistical comparison, 1 out of 5 (i.e. Never) is equivalent to 0% and 5 out of 5 (i.e. A lot) is equivalent to 100%. Thus, 2 out of 5 = 25%, 3 out of 5 = 50% and 4 out of 5 = 75% (see Appendix B.1 of [24] for the full results).

3.3 Overall probe/irrelevant recognition

The three Probe (familiar-lecturer) faces that were included in the experiment were reported to have been seen 33.9% of the scale (Mean confidence rating of 2.4 out of 5), and participants reported a very high (pre-experimental) familiarity of 94.6% with these Lecturer faces (4.8 out of 5). When comparing this to the (absent) Probe faces that were not included in the experiment, participants reported a similar high (pre-experimental) familiarity of 79.2% (4.2 out of 5), and only reported seeing these 'absent' familiar-lecturers 3.6% of the scale (1.1 out of 5), which is approximately one-ninth of the Lecturers that were, indeed, included in the experiment.

The three Irrelevants (*unknown*-lecturer) faces that were included in the experiment were reported to have been seen 4.8% of the scale (1.2 out of 5), and, similarly, the absent Irrelevant faces that were not included in the experiment were reported to have been seen 2.4% of the scale (1.1 out of 5). Finally, participants reported an imperceptible (pre-experimental) familiarity of 0% with all the Irrelevant/*unknown*-lecturer faces (1.0 out of 5).

As we were comparing Probe faces with Irrelevant faces, it was encouraging to discover that Probes were reported to have been seen 33.9% of the scale ($M = 2.4$; $SD = 1.2504$), which was seven times more than Irrelevants that were reported 4.8% of the scale ($M = 1.2$; $SD = 0.428$). Note that both conditions (Probes and Irrelevants) were, in fact, presented an equal number of times. The mean confidence rating of the main comparison conditions, for all participants, reveals a highly significant difference between the Probe (known-lecturer) faces and the Irrelevant (unknown-lecturer) faces, using pair-wise comparison ($M = 1.1714$, $SD = 1.5122$), $t(13) = 2.898$, $p = 0.0125$, $d = 1.2545$).

3.4 EEG analyses

We were interested in the EEG data across all the midline electrodes (Pz, Cz and Fz), but in-line with [30], we expect the strongest brain responses to familiar faces, to be recorded at the Pz electrode. Therefore, we will start by focusing on the Pz electrode, reporting Time and Frequency domain analyses (at group and participant level), before reporting the same analyses at Fz and Cz.

3.4.1 Pz electrode

At group-level, the grand average ERPs of all three critical stimuli (i.e. Target, Irrelevant and Probe), at the Pz electrode, revealed a clear difference between conditions (see Fig. 2). The Target was task-relevant, so it elicited a large classic P3. The Irrelevant (an unknown face), did not present a similar pattern to the Probe or Target. This was as expected, because non-salient stimuli were unlikely to breakthrough into conscious awareness, due to the high presentation rate of the RSVP streams. Although, the Irrelevant exhibits an interesting negative deflection, peaking at 300 ms, which we discuss later. In

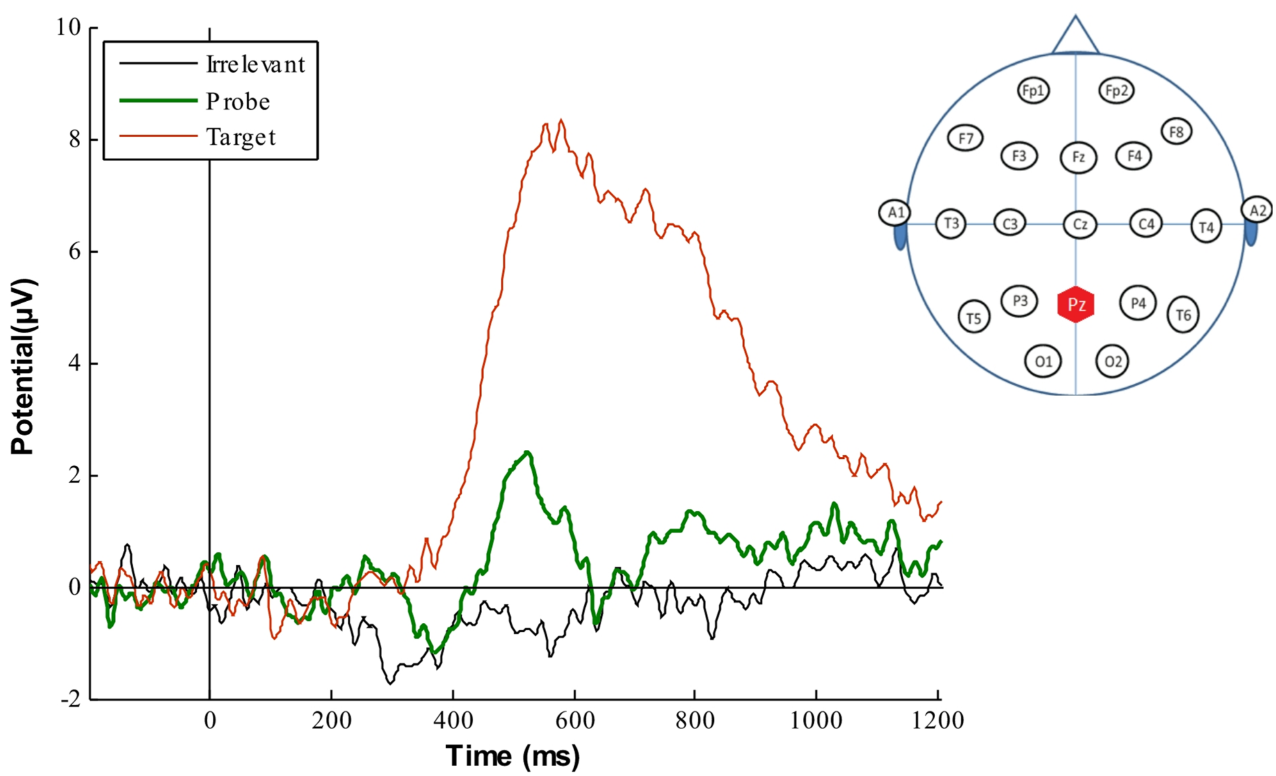

Fig. 2 Grand average ERPs elicited by the three critical stimuli (Irrelevant, Probe and Target), at Pz, showing a classic P3 pattern for the Target (in red, peaking at +8.5 μV), an oscillatory pattern for the Probe (in green, with an observed frequency of approx. 3–4 Hz) along with a late positivity, and a different pattern for the Irrelevant (in black, containing an interesting negative deflection, peaking at 300 ms) that is distinct from the Probe and Target. As before, Target was the stimulus that the participant was instructed to look for, whereas, they were not informed of the presence of the Probe (familiar lecturer face). The oscillatory pattern for the Probe suggests a significant difference with the Irrelevants (unknown lecturer faces)

a similar fashion to the famous faces experiment [1], the Probe (familiar face) exhibited evidence of a continuous oscillatory pattern, within a 280 to 620 ms time frame (and observed frequency of approximately 3–4 Hz) and a late positivity.

3.4.2 Group-level analysis, at Pz

Within the a-priori P600f time-frame, the AGAT method independently identified an orthogonal contrast 100 ms time window, at 453 to 553 (see Fig. 3), and our statistical tests produced a highly significant difference between the Probe ($M = 2.0086$, $SD = 2.0809$) and Irrelevant ($M = -0.4146$, $SD = 1.3911$), at the Pz electrode ($M = 2.4232$, $SD = 2.7311$), $t(13) = 3.3198$, $p = 0.0055$, $d = 1.3691$.

Unlike the famous faces experiment [1], the same statistical tests on the N400f component did not result in a significant difference between the two conditions: ($M = -0.57$, $SD = 1.1205$), $t(13) = -1.9035$, $p = 0.0794$, $d = 0.4532$. Furthermore, participant-level statistical tests (i.e. Monte Carlo permutation) on the N400f component confirmed that none of the 14 participants (0%) showed critical-significance (0.05 alpha level) between Probe and Irrelevant (Mean p-value = 0.5744, $SD = 0.2634$). As can be seen in Fig. 3, the unusual negativity in the Irrelevant condition (peaking at approximately 300 ms, and overlapping the Probe condition's N400f component) meant that the results of our statistical tests on N400f were not significant. In the remainder of the paper, we focus on the analysis of the P600f component.

3.4.3 Per-participant analysis

We performed statistical analyses of the ERP data, to determine whether the elicited response by the Probe (i.e. familiar lecturer face) was significantly different from that elicited by the Irrelevant (i.e. unknown lecturer face), on a per-participant basis. As shown in Table 1, per-individual statistical tests of Pz electrode's P600f component, resulted in 8 of 14 participants (57.1%) achieving significance (at alpha level $p < 0.05$, shown in green) for the Probe—Irrelevant comparison.

As shown in Fig. 4, the majority of participants' Probes elicited a clear positive deflection, within the aERPt identified highest positive 100 ms time window (highlighted in yellow), of the P600f search window (300 to 900 ms, based on prior-precedent from literature). However, relative to the Irrelevant, the Probe for five participants (nos. 1, 3, 4, 6 and 12) failed to show a significant positivity for P600f, and one participant (no. 8) failed because the Probe ERP was consistently below the Irrelevant ERP.

3.4.4 time frequency analysis (TFA)

As outlined in *Frequency Domain Analysis* (see Analysis section, above), to analyse the power and coherence of the EEG data, we have employed two Time Frequency transforms: Even-Related Spectral Perturbation (ERSP) and Inter-Trial Coherence (ITC), using EEGLAB's toolbox [21]. Whereas ERSP reflects the extent to which the signal power changes relative to the baselining window, ITC reflects the phase consistency (or synchronisation) between trials. ERSP/ITC changes in coherence enables us to quantify the multi-cycle oscillations that we had observed in the ERPs.

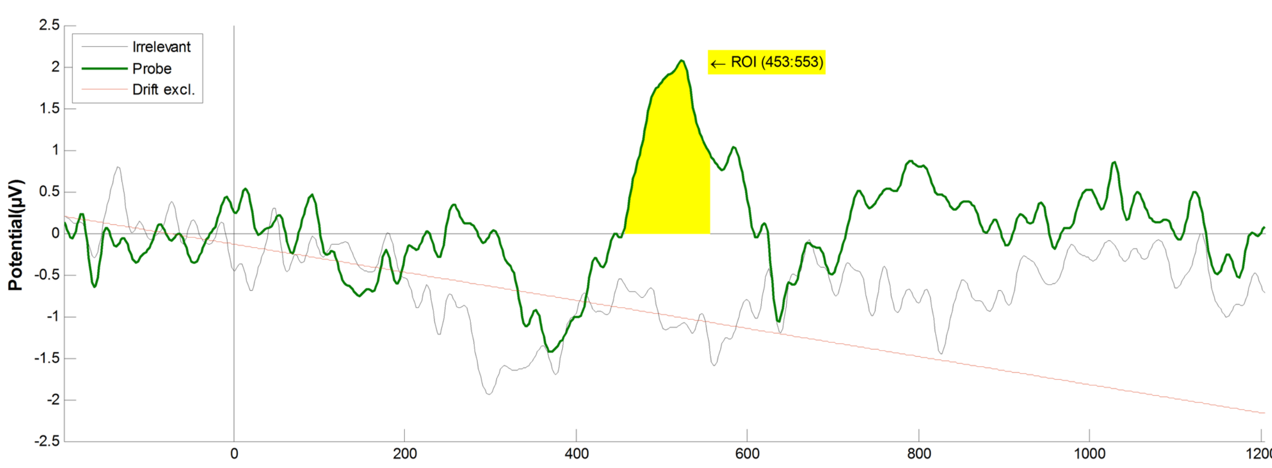

Fig. 3 Grand average ERPs elicited by Probe and Irrelevant (i.e. the main comparison conditions) at Pz, showing an oscillatory pattern for the Probe condition (in green), which does not exist for the Irrelevant condition (in black). The linear Drift has been excluded (in red, at −9.7% to the vertical), with a detrending method. The window identified by the AGAT method is shown in yellow. Even though participants were not informed of the presence of the Probe (familiar lecturer face), statistical tests show a highly significant difference between Probe and Irrelevant, for P600f ($t(13) = 3.3198$, $p = 0.0055$, $d = 1.3691$). However, the same statistical test on N400f was not significant ($t(13) = -1.9035$, $p = 0.0794$, $d = 0.4532$)

Table 1 Per-participant analysis, at Pz, for the P600f component, showing the mean amplitude values of the Probe and Irrelevant conditions, from the same 100 ms time window, which was independently found using the aERPt method

Participant	Probe (*M*)	Irrelevant (*M*)	*p*-value
1	− 0.8440	− 2.4415	0.166
2	3.4664	0.1121	0.036
3	0.7447	1.5237	0.651
4	2.0746	1.7694	0.435
5	0.5265	− 1.9373	0.048
6	1.9794	0.7962	0.328
7	4.2431	− 0.1308	0.002
8	− 1.1190	0.5188	0.858
9	1.4428	− 1.2087	0.008
10	2.0194	− 1.3143	0.027
11	4.6768	− 1.5233	< 0.001
12	0.1841	0.5206	0.601
13	2.6896	− 0.1072	0.034
14	6.0366	− 2.3814	< 0.001

Significant participants are shown in green. Out of 14 participants, eight (57.1%) achieved critical-significance. p-values are derived from a Monte-Carlo permutation test analogue of a two-sample independent t-test, involving 1000 resamplings

We applied a notch filter, between 7 and 9 Hz, in order to filter out any Steady State Visual Evoked Potential (SSVEP) set up by the RSVP stream. Thus, as long as there are no significant power increases at higher frequencies, we could fix the upper boundary of our analysis at 7 Hz (also noting that the lowest boundary is fixed by our standard high-pass filter, on 0.5 Hz). Even so, in addition to the fixed-boundary analysis window (0.5 to 7 Hz), we also performed the full ERSP/ITC analyses on the full frequency range (0.5 to 45 Hz), to assess the power/coherence changes at higher frequencies.

3.4.5 Group-level TFA

As seen in Fig. 5, ERSP and ITC results of the AGAT of the Probe and Irrelevant conditions at Pz, exhibits a large power, respectively coherence, increase around 300 to 650 ms (post-stimulus), mainly, at the low frequencies. To compare Probe and Irrelevant at the group-level, two measures were obtained for each participant, and a two-tailed paired *t*-test was used to perform statistical inference for ERSP and ITC. As can be seen in the grand-Probe versus grand-Irrelevant ERSP/ITC comparisons (see Fig. 6), increases in power/coherence are predominantly evident in the grand-Probe condition, which suggests detection of the familiar lecturer face (ERSP > 4 dB, and ITC > 0.4). However, the grand-Irrelevant condition lacks any significant power/coherence fluctuations, within the same time window, which implies little-to-no conscious or sub/liminal detection of the unknown lecturer face. Nearly all the power/coherence fluctuations are occurring in the lower frequency range (i.e. 0.5 to 7 Hz).

Over the full frequency-range (i.e. 0.5 to 45 Hz), the group-level analysis at Pz for ERSP was significant, at the AGAT defined window 461 to 561 ms (see Fig. 5), confirming a difference between the Probe and Irrelevant conditions ($t(13) = 2.9737$, $p = 0.0108$, $d = 1.0417$). For the group-level ITC over the same/maximum frequency range, our statistical test was not significant, at the AGAT defined window 334 to 434 ms: ($t(13) = 1.8064$, $p = 0.0941$, $d = 0.7997$).

However, focusing on the narrower frequency-band (i.e. 0.5 to 7 Hz), the group-level analysis at the Pz electrode for ERSP was highly significant, at the AGAT defined window 484 to 584 ms (see Fig. 7), confirming a difference between Probe and Irrelevant ($t(13) = 3.4769$, $p = 0.0041$, $d = 1.2649$). For group-level ITC over the same (narrower) frequency range, our statistical test was also significant at the AGAT defined window 588 to 688 ($t(13) = 2.322$, $p = 0.0371$, $d = 0.7442$).

3.4.6 Per-participant-level TFA

Per-participant statistical inference—in the form of a randomisation test—confirmed the substantial increase in the Probe's power (ERSP) and coherence (ITC), as compared to the Irrelevant. Statistical tests of the *narrower*

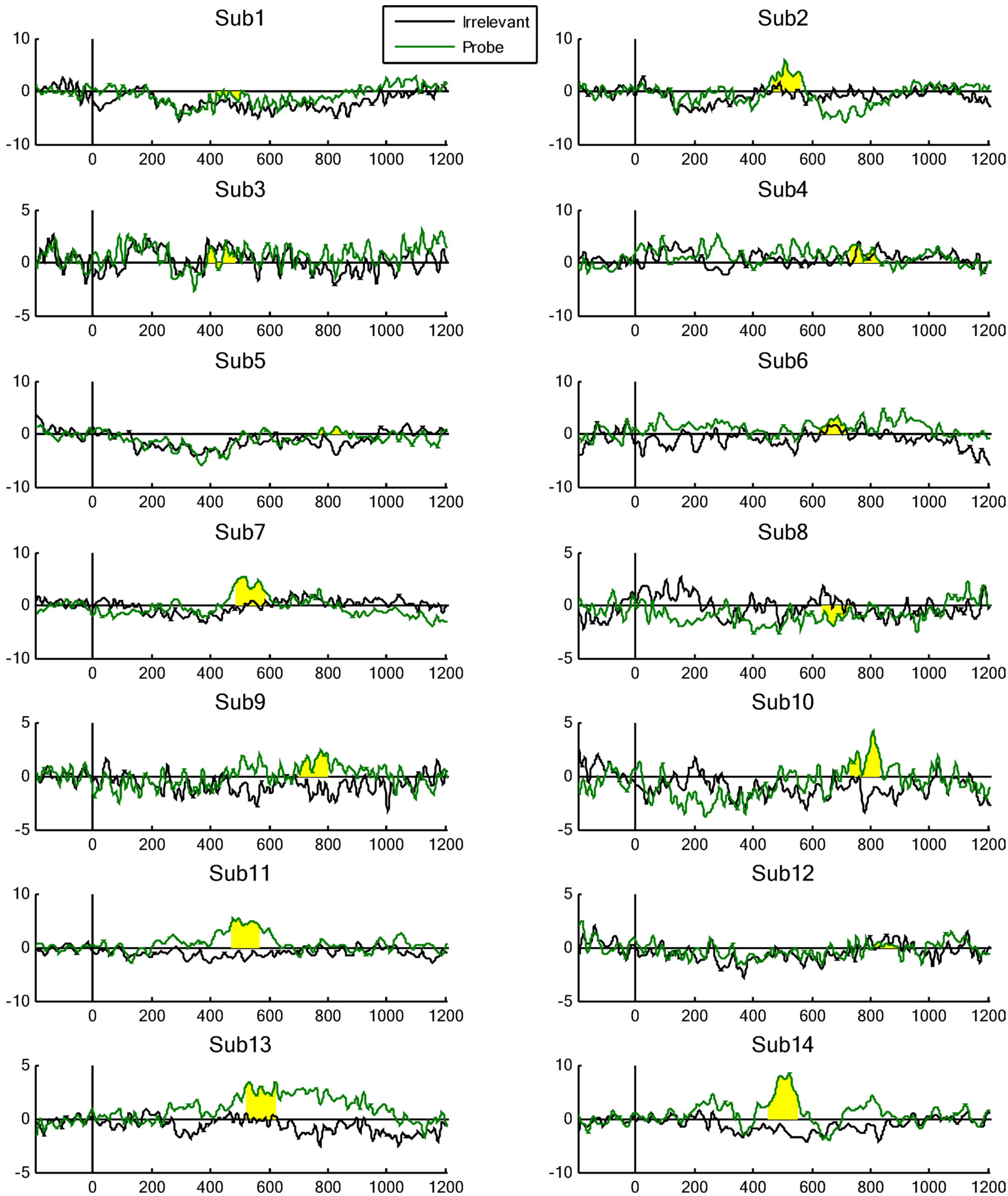

Fig. 4 Per-participant Probe (in green) and Irrelevant (in black) ERPs, at the Pz electrode (x-axis represents Time in milliseconds, and y-axis represents Potential in microvolts). Each ERP shows the orthogonally identified highest positive 100 ms time window (yellow highlight) for P600f (using the aERPt method), where 8 of 14 participants (57.1%) show a significant difference between the Probe and Irrelevant conditions. Note, y-axis scales vary across subplots

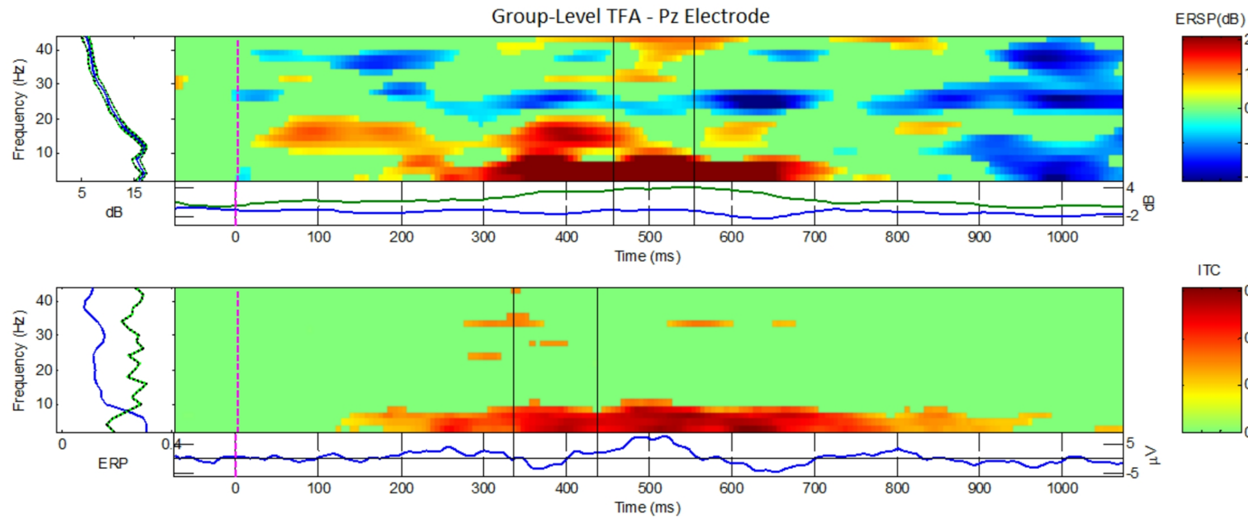

Fig. 5 Group-level Time Frequency plots, at the Pz electrode, using the combined Probe and Irrelevant conditions, i.e. the AGAT. The top plot depicts the ERSP, and the bottom plot depicts the ITC. The independent window selection (using AGAT method) for ERSP produced a Region of Interest (ROI) at 461:561 ms, and earlier ROI for ITC, at 334:434 ms. Increases in power/coherence have been mostly concentrated in the 0.5 to 10 Hz frequency range, and are strongest in the ROI time frames (SSVEP has been filtered out, by applying a 7:9 Hz notch filter)

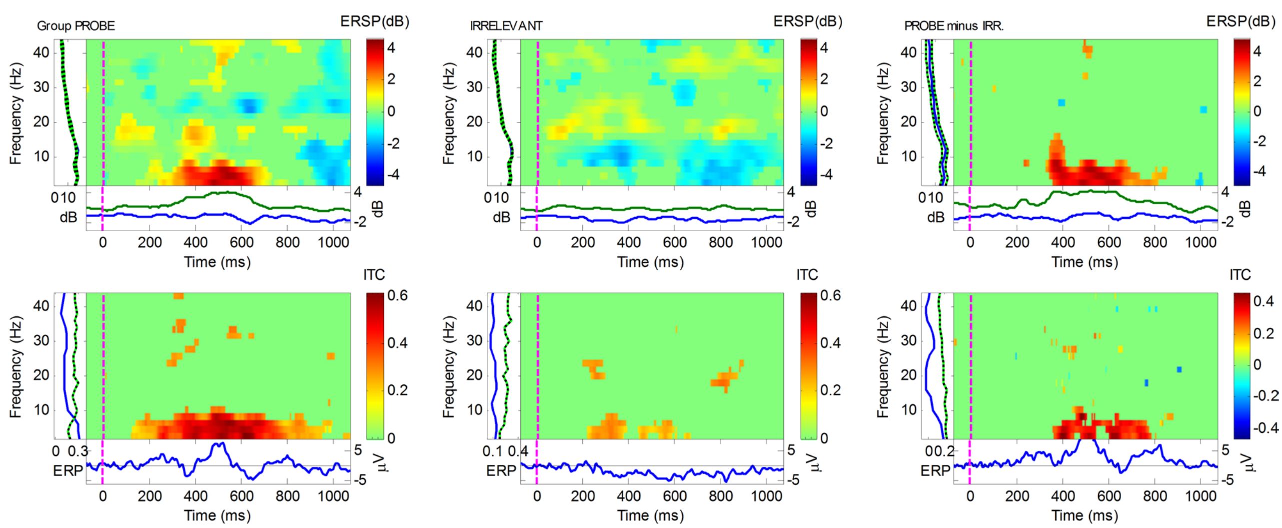

Fig. 6 Group-level Time Frequency Analysis, at Pz electrode, for the difference between critical stimuli (Probe and Irrelevant), across the full frequency range (0.5 to 45 Hz), showing ERSP (top row) and ITC (bottom row). The first column of ERSP/ITC plots show the power/coherence changes in the grand-Probe condition, and the second column shows the same for the grand-Irrelevant condition. The third column is the difference between grand-Probe and grand-Irrelevant (i.e. Probe minus Irrelevant), which confirms group-level increases in power/coherence for the grand-Probe only: ERSP ($t(13) = 2.9737$, $p = 0.0108$, $d = 1.0417$), and ITC ($t(13) = 1.8064$, $p = 0.0941$, $d = 0.7997$). Note that at each frequency and time point, increases in power/coherence are in red; decreases in blue, and green indicates no significant change

frequency band (0.5 to 7 Hz), resulted in two orthogonal contrast time windows for each participant (ERSP average window: 496 to 569, and ITC average window: 413 to 513) and *p*-values that revealed a difference between the Probe and Irrelevant conditions (see Table 2). For ERSP, 10 out of 14 participants' *p*-values (71.4%) were significant. Similarly, for ITC, 9 out of 14 participants' *p*-values (64.3%) were significant.

According to the above Frequency domain analysis (see Table 2), with the exception of participant 1, statistical test results of ITC appear to be closely correlated to the Time domain's statistical tests of the ERP data (see Table 1), at per-participant-level, since both analyses produce significant *p*-values for participants 2, 5, 7, 9, 10, 11, 13 and 14.

Finally, even though we have justified fixing the upper boundary of our analysis at 7 Hz (i.e. due to SSVEP waveform, which required a notch-filter on 7 to 9 Hz), we confirmed that per-participant statistical analysis of the maximum frequency range (0.5 to 45 Hz), resulted in

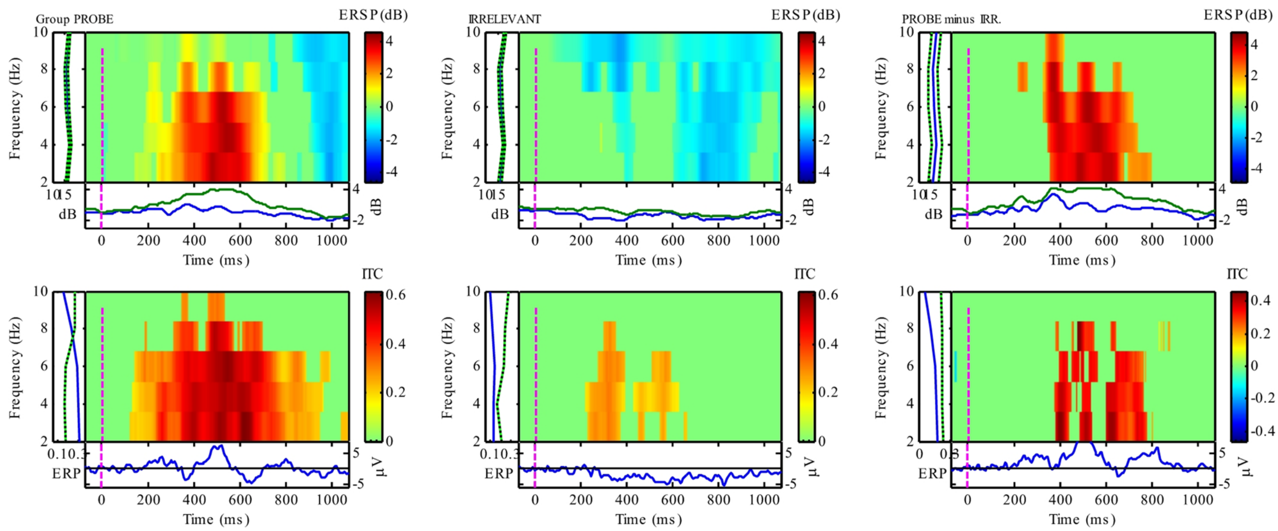

Fig. 7 Group-level Time Frequency Analysis, at Pz, for the difference between critical stimuli (Probe and Irrelevant), at the narrower frequency band (0.5 to 7 Hz), showing ERSP (top row) and ITC (bottom row). The first column of ERSP/ITC plots show the power/coherence changes in the grand-Probe, and the second column shows the same for the grand-Irrelevant. The third column is the difference between grand-Probe and grand-Irrelevant (i.e. Probe minus Irrelevant), which confirms group-level increases in power/coherence for the grand-Probe only: ERSP ($t(13) = 3.4769$, $p = 0.0041$, $d = 1.2649$), and ITC ($t(13) = 2.322$, $p = 0.0371$, $d = 0.7442$). At each frequency and time point, increases in power/coherence are in red; decreases in blue, and green indicates no significant change

Table 2 Per-participant time frequency analysis of power (ERSP) and coherence (ITC), at Pz electrode, using the narrower frequency range (0.5 to 7 Hz).

Participant	ERSP *p*-values	aERPt win.	ITC *p*-values	ITC win.
1	0.6733	703	0.0107	369
2	< 0.0001	508	< 0.0001	525
3	0.04	461	0.2493	162
4	< 0.0001	404	0.22	709
5	0.008	410	0.0147	520
6	0.0147	531	0.6213	260
7	< 0.0001	514	0.0293	428
8	0.2147	795	0.244	53
9	0.2907	473	0.0187	674
10	0.264	307	0.043	629
11	< 0.0001	490	< 0.0001	484
12	0.0253	559	0.284	53
13	0.0133	455	0.04	473
14	< 0.0001	334	< 0.0001	449

For each participant, an orthogonal contrast time window was employed (using the aERPt method), and p-values were obtained for ERSP and ITC, by comparing the Probe and Irrelevant conditions, using a randomisation test. 10 of 14 ERSP p-values (71.4%) were significant, and 9 of 14 ITC p-values (64.3%) were significant. p-values are derived from a Monte-Carlo permutation test analogue of a two-sample independent t-test, involving 10,000 resamplings

p-values that revealed a difference between the Probe and Irrelevant conditions (see Table 3). For ERSP, 7 out of 14 participants' *p*-values (50%) were significant. As for ITC, 9 out of 14 participants' *p*-values (64.3%) were significant.

3.4.7 Other midline electrode sites

All the above Time and Frequency domain analyses focused on the Pz electrode, but we were also interested in the other two midline electrodes (Cz and Fz), to confirm that, in line with [30], the strongest brain responses to familiar faces are recorded at Pz. The following analogous Time domain analyses of Fz and Cz, aim to find out if the P600f evoked by the Probe was significantly different from that evoked by the Irrelevant.

3.4.7.1 Fz electrode At the group-level, the grand-average ERPs of the two critical stimuli (i.e. the Probe and Irrelevant), at the Fz electrode, revealed a clear difference between the conditions (see Fig. 8). The Irrelevant did not present any feature/pattern similar to the Probe. Again, the Probe elicited a continuous oscillatory pattern, within a 200 to 620 ms time frame. This waveform at Fz is simi-

Table 3 Per-participant time frequency analysis of power (ERSP) and coherence (ITC), at Pz electrode, using the maximal frequency range (0.5 to 45 Hz)

Participant	ERSP *p*-values	aERPt win.	ITC *p*-values	ITC win.
1	0.35	455	0.002	357
2	0.002	484	0.001	381
3	0.056	738	0.093	191
4	0.004	422	0.04	53
5	0.006	307	0.136	410
6	0.038	53	0.025	795
7	0.001	525	0.012	455
8	0.157	756	0.166	490
9	0.391	59	0.002	635
10	0.183	600	0.002	53
11	0.105	289	0.092	145
12	0.06	600	0.332	53
13	0.043	53	0.05	496
14	< 0.001	461	< 0.001	334

For each participant, an orthogonal contrast time window was employed (using the aERPt method), and p-values were obtained for ERSP and ITC, by comparing the Probe and Irrelevant conditions, using a randomisation test. 7 of 14 ERSP p-values (50%) were significant, and 9 of 14 ITC p-values (64.3%) were significant. p-values are derived from a Monte-Carlo permutation test analogue of a two-sample independent t-test, involving 1000 resamplings

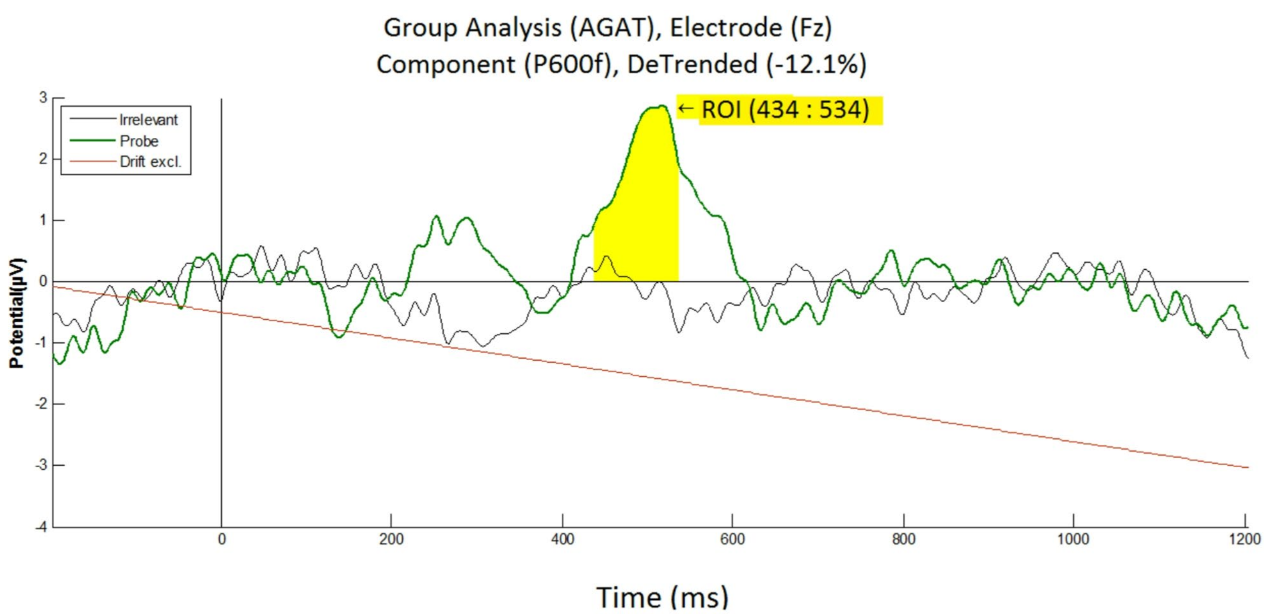

Fig. 8 Grand average ERPs elicited by the Probe and Irrelevant, at Fz, showing an oscillatory pattern for the Probe (in green), which does not exist for the Irrelevant (in black). The linear Drift has been excluded (in red, at – 12.1% to the vertical) with a detrending method. Baseline correction is between – 100 ms to zero. Even though participants were not informed of the presence of the Probe, statistical tests show a significant difference between Probe and Irrelevant for P600f ($t(13) = 3.3151$, $p = 0.0056$, $d = 0.9834$)

lar to the oscillatory waveform at Pz (see Fig. 3), and it confirms a large difference between the Probe (familiar lecturer face) and the Irrelevant (unknown lecturer face), at frontal midline electrodes. At this Fz electrode, an orthogonal contrast time window, for the highest positive (P600f) component, was found (using the AGAT method) at 434 to 534 ms.

Statistical analyses at Fz—a paired *t*-test of the mean amplitudes of Probe and Irrelevant, across all participants—was run at the group level for the P600f component. Our statistical tests produced a significant difference between the Probe ($M = 2.0696$, $SD = 2.4934$) and Irrelevant ($M = -0.0349$, $SD = 1.7156$), at Fz: ($M = 2.1045$, $SD = 2.3754$), $t(13) = 3.3151$, $p = 0.0056$, $d = 0.9834$.

At the Fz electrode, per-participant statistical tests (i.e. Monte Carlo permutation) on the P600f component confirmed that 6 of 14 participants (42.9%) showed significant differences between Probe and Irrelevant (see Table 4). This confirms that results at Pz (i.e. 8 of 14;

Table 4 Per-participant analysis, at Fz, for the P600f component, showing the mean amplitude values of the Probe and Irrelevant conditions, from a 100 ms time window, found using the aERPt method

Participant	Probe (*M*)	Irrelevant (*M*)	*p*-value
1	− 2.2370	− 1.9522	0.590
2	4.5957	− 1.0226	0.007
3	2.7125	2.4847	0.463
4	5.8049	1.4776	0.012
5	3.9095	0.2096	0.020
6	2.1781	1.2152	0.329
7	6.2001	− 0.9125	< 0.001
8	− 0.8330	1.4425	0.909
9	0.9747	− 0.8974	0.079
10	1.3183	0.0575	0.288
11	4.3378	0.7975	0.016
12	0.6016	0.6258	0.533
13	5.5821	1.8344	0.016
14	1.8365	− 0.8767	0.061

Statistical tests on P600f resulted in 6 of 14 participants (42.9%) being significant, which is not as high as equivalent results at Pz (i.e. 8 of 14: 57.1%), confirming that Fz failed to show a stronger brain response when compared to Pz. p-values are derived from a Monte-Carlo permutation test analogue of a two-sample independent t-test, involving 1000 resamplings

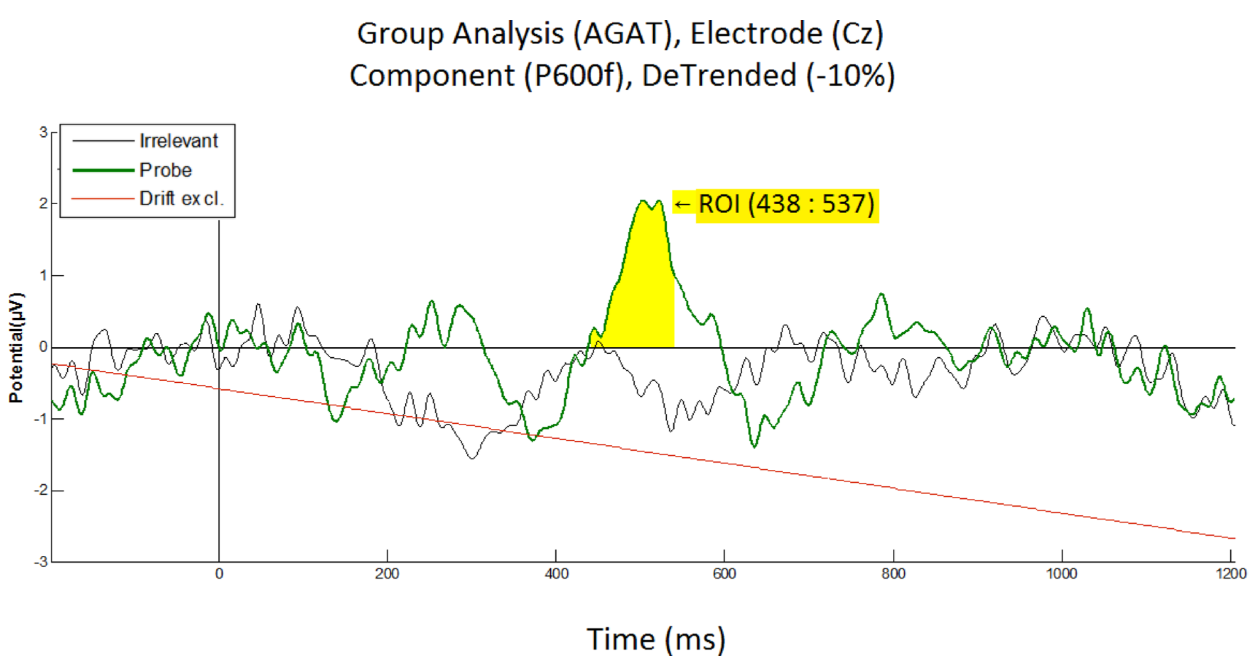

Fig. 9 Grand average ERPs elicited by the Probe and Irrelevant, at the Cz electrode, showing an oscillatory pattern for the Probe condition (in green), for the Irrelevant condition (in black). The linear Drift has been excluded (in red, at – 10% to the vertical) with a detrending method. Even though participants were not informed of the presence of the Probe (familiar lecturer face), a significant difference between Probe and Irrelevant, for P600f was observed ($t(13) = 2.8346$, $p = 0.0141$, $d = 0.9649$)

57.1%—see Table 1), were better than at Fz, agreeing with studies [30] that report stronger brain responses (to familiar faces) at Pz.

3.4.7.2 Cz electrode The same group-level analysis that was carried out at Fz, was performed at Cz, revealing a difference between the Probe and Irrelevant conditions (see Fig. 9). Once again, the Probe elicited a continuous oscillatory pattern, within a 220 to 720 ms time frame, which was not present for the Irrelevant. This waveform at Cz is very similar to the oscillatory waveforms at Pz and Fz (see Figs. 3 and 8, respectively). At this Cz electrode, an orthogonal contrast time window, for the highest positive (P600f) component, was found (using the AGAT method) at 438 to 537 ms. A paired t-test for the P600f produced a significant difference between the Probe ($M = 1.2271$, $SD = 2.0023$) and Irrelevant ($M = -0.4156$, $SD = 1.3367$), at Cz: ($M = 1.6427$, $SD = 2.1683$), $t(13) = 2.8346$, $p = 0.0141$, $d = 0.96494$.

At the Cz electrode, per-participant statistical tests (i.e. Monte Carlo permutation) on the P600f component confirmed that 4 of 14 participants (28.6%) showed a significant difference between Probe and Irrelevant. Our per-participant results (see Table 5), confirmed that the effect was stronger at Pz (i.e. 8 of 14; 57.1%—see Table 1) than at Cz; once again, agreeing with studies that report stronger brain responses (to familiar faces) at the Pz electrode.

Finally, we have demonstrated that all three midline electrodes (Pz, Fz and Cz) have exhibited similar oscillatory waveforms, and that statistical tests showed significant differences between Probe and Irrelevant. Although our choice to focus on the Pz electrode was a priori (in-line with [30]), we found that the strongest brain responses to familiar lecturer faces, was indeed recorded at the Pz electrode, as we observed in [1].

4 Discussion

The primary aim of the experiment presented here was to investigate whether faces that are personally known to an individual (familiar lecturer faces) can breakthrough into conscious awareness, and that event can be detected by EEG, on a group and per-participant basis, as successfully as our previous famous faces experiment [1]. To do this, we performed statistical inference on ERP data (in the Time domain) and single-trial data (in the Frequency domain), to determine whether the evoked response by the Probe (familiar lecturer) faces were significantly different from that evoked by the Irrelevant (unknown lecturer) faces.

The key comparison was between Probe faces and Irrelevant faces, for which a significant difference was observed between the ERPs, at all three mid-line electrodes (Pz, Fz and Cz); see [24] for more details. However, unlike the prominent negative-and-positive deflections for the Probe (i.e. N400f, followed by P600f) we previously observed for famous faces [1], this experiment's negative deflection (N400f) was muted, but its positivity (P600f) was equivalent (over a similar time frame of 300 ms to 600 ms). Additionally, even though there was again evidence of an oscillatory pattern, this second experiment's ERPs are somewhat different and are typically weaker than those of the earlier famous faces experiment.

We closely mirrored the famous faces experiment's design and analysis, with the main change being to the critical stimuli (i.e. replacement of the *celebrity* faces with *lecturer* faces). Even though the first (celebrity faces)

Table 5 Per-participant analysis, at Cz electrode, for the P600f component, showing the mean amplitude values of the Probe and Irrelevant, from the same 100 ms time window, found using the aERPt method

Participant	Probe (M)	Irrelevant (M)	p-value
1	−1.1798	−2.9437	0.187
2	3.4261	−0.1361	0.027
3	1.6282	1.6906	0.504
4	0.6658	1.8322	0.726
5	0.3553	−0.2447	0.351
6	2.9873	0.8060	0.210
7	4.4588	−0.2104	0.001
8	−0.8165	1.1491	0.879
9	0.9685	−1.0082	0.062
10	1.2720	−0.2308	0.199
11	4.2569	−0.0883	0.002
12	0.4085	0.7474	0.602
13	3.5615	0.7885	0.061
14	3.4331	−1.8483	0.002

Statistical tests on P600f resulted in 4 of 14 participants (28.6%) being significant, which is not as high as equivalent results at Pz (i.e. 8 of 14: 57.1%). p-values are derived from a Monte-Carlo permutation test analogue of a two-sample independent t-test, involving 1000 resamplings experiment [1] established the viability of using faces to infer recognition in the RSVP paradigm, the next logical step was to investigate whether faces that are personally known to an individual would have a similar effect. After all, deception detection tests would seek to demonstrate a person under investigation's recognition of somebody he/she has a personal relationship/familiarity with, rather than the non-partisan knowledge of a famous person.

In keeping with the famous faces experiment, participants were not informed that familiar faces may appear in the RSVP streams, and yet, our statistical tests confirmed the breakthrough events. Having been instructed to only look for the Target face, the inclusion of Probe faces (i.e. University lecturers) assessed participants' ability to perceive incidental intrinsically salient faces. Statistically comparing brain responses to Probes and Irrelevants, in the time, as well as, frequency domains, enabled us to confirm differential perception of the Probe over the Irrelevants (i.e. unknown faces), at group and participant levels.

4.1 Time domain

At the Pz electrode, group-level analysis of ERPs confirmed the significance of the difference between the grand-Probe and grand-Irrelevant ($p = 0.0055$), and per-participant statistical analyses of ERPs confirmed that a total of 8 of 14 participants (57.1%) had significant *p*-values, revealing a difference between Probe and Irrelevant. Again, we confirmed that the brain response to the Probe at the per-participant level (8 of 14), achieved more participants at the Pz electrode, than Fz ($p = 0.0056$, and 6 of 14) or Cz ($p = 0.0141$, and 4 of 14). (A multiple comparisons correction across all group-level tests in this paper did not lead to the loss of any significant results; see Appendix 2.)

4.2 Frequency domain

At the Pz electrode, group-level Time Frequency analysis, across the narrower frequency band (0.5 to 7 Hz), confirmed the significance of the difference between the grand-Probe and grand-Irrelevant for ERSP ($p < 0.0041$) and for ITC ($p = 0.0371$). Per-participant statistical inference, across the same frequency band (0.5 to 7 Hz), confirmed that 10 out of 14 participants' *p*-values (71.4%) were significant. As for ITC, per-participant statistical inference showed that 9 out of 14 participants' *p*-values (64.3%) were significant, confirming the difference between the Probe and Irrelevant conditions. (A multiple comparisons correction across all group-level tests in this paper did not lead to the loss of any significant results; see Appendix 2.)

Thus, our experimental findings show substantial differences between the Probe and Irrelevant, which we believe result from the former stimuli reaching conscious awareness and generating pronounced electrical responses (as seen in the Probe ERPs), as suggested by our recognition memory experiments. However, there was an interesting new electrical response: a negative deflection to Irrelevant, peaking at 300 ms, which we did not observe in the famous faces experiment (see for example, black line in Fig. 2).

It is noteworthy that participants did not report seeing the Irrelevant (unknown lecturer) faces, i.e. their recognition memory performance for Irrelevants was very poor. Accordingly, we have formed two theories to potentially explain this phenomenon. The first is that this posterior negativity may be related to subliminal registering (i.e. a covert response) of a repetition by the brain, and the second is that it may relate to an incongruity, between the Irrelevant and filler/distractor images, which was not present in the previous (famous faces) experiment, or indeed, both. More specifically, the Irrelevant images in the previous (famous faces) experiment were chosen randomly from the Distractor database, but the Irrelevant images in this (lecturer faces) experiment did not come from the Distractor database—they were (unknown) lecturer faces from a different University. In particular, to ensure a fair comparison with Probes, we used faces of academics that were unfamiliar to the participants. This necessarily, meant a "category" change between Distractors and Irrelevants. The key aspect of this change may just relate to changes in low-level visual features, since the Irrelevants were photos taken by the experimenter.

It also seems likely that this difference did not generate a breakthrough event for Irrelevants, since, as previously discussed, end of experiment, memory performance for Irrelevants was low. Finally, and importantly, while the seemingly subliminal nature of this evoked response to the Irrelevants means it could be of scientific interest, the presence of such a response does not confound the logic of our experiment, although it may weaken the differences between Probe and Irrelevant that we observe.

4.3 Concealed information test

Importantly, although in the work presented here participants were not explicitly instructed to lie, the instruction and experiment set-up was consistent with how the concealed information test would unfold, given our previous experimental work. Specifically, participants were instructed to search for a target face. In the classic RSVP-CIT paradigm [12, 13], this would be the Fake—which participants are instructed to search for in RSVP streams. The Fake/Target does not get analysed in either paradigm but rather is present to give participants a task and thereby, hold their attention on the streams. In this paper, the Probe stimulus arose as it could do in a CIT, specifically, there was no instruction indicating that such a stimulus would be present and its intrinsic salience/familiarity caused it to break into awareness.

Our work on countermeasures-resistance [12], suggests that participants cannot subvert the test by either attempting not to see the Probe or attempting to see an Irrelevant. This finding was made with name stimuli [12]. Thus, it would be valuable to confirm this finding with the face stimuli used in this paper. However, from our theoretical work [3, 7–9], we believe that fundamental characteristics of presentation of visual stimuli in RSVP underlie this counter-measures resistance, which one would expect to be stimulus-type independent. Accordingly, this resistance to counter-measures suggests that the key demonstration of the Fringe-P3 method's validity with a new stimulus type, is what we have performed in this article. That is, the simple question to answer is whether familiar faces "pop-out" of streams at a presentation rate where non-familiar stimuli do not. This is the question we have explored here for familiar faces.

4.4 Size of effect

The per-participant statistical results were not as strong as those for the famous faces experiment. This is likely to be because the familiar (lecturer) faces do not have the level of salience of the famous faces. However, it may also be impacted by the previous exposure to the *images* of the famous *celebrity* faces (i.e. published photographs, which were frequently in the public eye, and surely seen by many participants, on many occasions and over a long period of time). In contrast, the *lecturer* faces of the experiment presented here, whilst having some personal salience to the participants, their *photographs* were seen for the first time, in the fast moving RSVP stream of images. In fact, the behavioural/recognition tests for both experiments support this premise, as participants in the previous famous faces experiment reported seeing the (celebrity) Probes with an average confidence rating of 3.4 out of 5 (60%), whereas, participants in the current experiment reported seeing the (lecturer) Probes with an average confidence rating of only 2.4 out of 5 (33.9%).

One possible explanation for weaker findings is that participants did not "see" all three of the Probe faces presented in the experiment. This may be because they did not have a strong familiarity with all three of the selected Probe faces. Consequently, an interesting line for future work is to develop a statistically valid method to "close-in" during the experiment on the Probe face for which the participant has the most familiarity. Of course, simply selecting the largest brain response post hoc would inflate the type 1 error rate. As a result, some consideration of the appropriate statistical procedure, probably through ground-truth synthetic simulations, is required to take such a closing-in procedure forward.

In addition, future studies could take one step closer to the real-world scenario of concealed information tests, whereby, the hidden nature of the Probe critical stimuli would be to some extent revealed to the participants, at the start of the experiment. Having demonstrated that the breakthrough of familiar faces can be detected even when participants were not expecting to see celebrity-or-lecturer faces—because the presence of the Probe was concealed and the Target was the only task-relevant objective—we note that one could inform participants that, in addition to the Target (which will remain task-relevant), they may see a familiar lecturer face, in the RSVP streams. Naturally, one must not inform participants which lecture faces may appear in the experiment (or show them any photographs), but one can place their minds in a state where they are expecting to see the class of stimuli that the Probes will satisfy. This arrangement is closer to the real-life application of a deception detection test, in which the perpetrator would be fully aware of the purpose of the experiment (i.e. to find out if s/he is familiar with an accomplice). This strategic priming of the participants is likely to increase the effectiveness of the method.

4.5 Countermeasures and practicality

As previously emphasized, a major benefit arising from the fast presentation rate of the Fringe-P3 method is the demonstrated robustness against key countermeasures [12]. However, two other issues that might be considered a problem for the Fringe-P3 method are, 1) attending away and 2) individual-specific perceptual thresholds. We discuss these in turn.

1. *Attending away*: participants would affect the results of the method, if they avoid (or fail to observe) images that are presented in the middle of the screen. They could do this by shifting their attentional focus away from the stream, with covert shifts of attention potentially being particularly problematic, since they would not be detectable with eye tracking. However, if participants are attending to the stream, a Steady State Visually Evoked Potential (SSVEP) will be observable at electrodes over visual cortex and since one knows the frequency of the stream, this steady state response can be efficiently detected. Thus, absence of an SSVEP would indicate noncompliance. Additionally, one could use the Target image Hit/False Alarm rate, as a further indicator of attentional engagement.
2. *Individual-specific perceptual thresholds*: we expect a difference between the Probe (known face) and Irrelevant (unknown face), but in order to reduce opportunities for countermeasures, the stream's presentation rate must be high enough to only permit the Probe images to breakthrough into conscious

awareness. Indeed, Probes may not be perceived if the presentation speed is too high, and Irrelevants may breakthrough if the speed is too slow. However, there are considerable individual differences in perceptual thresholds, which suggest that the optimal presentation rate will vary across participants. Whilst we have adopted a fixed presentation rate (e.g. a-priori SOA of 133 ms), future studies could improve the breakthrough difference between the Probe and Irrelevant, by running a pre-experiment training session using a 'staircase' procedure, to determine the optimal presentation rate for each participant.

4.6 Potential confounds

An important confound that could arise in an experiment focussed on familiarity effects is if one selected stimuli that would be more familiar across the population, perhaps by selecting stimuli that were rated as more familiar, potentially by an independent group. These familiarity ratings might be biased towards stimuli that are more distinctive or affectively engaging. They are also of course likely to be more frequent in the environment.

Importantly, we did not select Probes to be intrinsically familiar stimuli in the world. We chose them with the following criteria: 1) they were lecturers in the Computer Science Department at the University of Kent; and 2) they had one or more PhD students. It is difficult to see how this selection would lead to Probe faces that were particularly distinctive or affectively engaging.

It is almost certain that the participants (graduate students) would have high exposure to the Probe faces. However, that is the very phenomenon that is likely to have made the face familiar to the participant in the first place. So, in and of itself, this is not a confound.

Importantly, we selected the Irrelevant images carefully to avoid confounds such as distinctiveness or how affectively engaging they are. Specifically, Irrelevants were also lecturer faces, but from a different University to that of the participants. Both Probe and Irrelevant images were taken with the same camera, from the same position/distance, and prepared identically.

Although participants will have less exposure to the Irrelevants than the Probes, which we believe is inherent to a face becoming more familiar.

4.7 Limitations

Study Sample: Importantly, our objective in this paper is proof of principle, i.e. to provide a first demonstration that a method such as we are proposing has the potential to be used with personally familiar faces. However, it is important to realise that to definitively characterize the effectiveness of the method in the broader population would require a further study with a broader, larger, cohort. While the demographics of our study are not untypical amongst academic work, they are limited in their population coverage (e.g. in respect of age and socio-economic profile). This said, in a different context, Hellerstedt, Moccia, Brunskill, Bowman & Bergström [27] observed that while the P3 is somewhat reduced in amplitude in an elderly cohort, its form is not greatly altered, suggesting findings from a young cohort do generalise.

Additionally, although our cohort is small in this study, just 14, and thus, we have few degrees of freedom, key group-level effects are large (e.g. time-domain P600f: $t(13) = 3.3198$, $p = 0.0055$, $d = 1.3691$; and frequency domain, narrow frequency band (0.5 to 7 Hz), ERSP: $t(13) = 3.4769$, $p = 0.0041$, $d = 1.2649$).

The point of uncertainty with a small data set like this is how to interpret non-significant findings (e.g. time-domain N400f: $t(13) = -1.9035$, $p = 0.0794$, $d = 0.4532$). These may well be false negatives and larger cohort studies are required to clarify these effects.

Due to the department in which the data was collected, the cohort was biased towards males. This should be rectified in future studies. However, notably, in other work, with a different gender balance (seven females and four males), we have observed similar findings with a different stimulus type (email addresses) [25].

Time-on-Task Effects: An important question is whether there are time-on-task effects, i.e. systematic changes across the blocks of our experiment. For example, there may be habituation effects on P3 amplitude or learning effects, such as the Irrelevant being recognised, due to its repeated presentation through the course of the experiment.

Unfortunately, effects across blocks are somewhat confounded in this data set, since the Probe stimulus was not controlled across blocks. In fact, we deliberately ordered the Probe stimuli from what we anticipated would be the most salient in the first block, to second most salient in the second, to least salient in the last. However, we have looked at these issues in previous work, across a range of stimulus types.

The following are key points concerning time-on-task effects:

1. We compared Irrelevant ERPs from early in the experiment to late in it in figure 5 in Bowman, Filetti, Janssen, Su, Alsufyani & Wyble [13]. There was no evidence that a P3 was present in either the early or late portions of the experiment, which stands against the possibility that the Irrelevant starts to be seen during the course of the experiment due to its frequent repetition.
2. In our countermeasures work, we actually informed participants of the presence of repeating Irrelevant stimuli and instructed them to look for these stimuli

[12]. As can be seen in Figs. 4 and 5 of that article, there was no evidence that these Irrelevants generated a P3. This absence of an effect of repetition on non-familiar items is consistent with our theoretical work on this topic [3, 7–9],
3. Notwithstanding the proviso concerning blocks not being controlled in the experiment reported in this paper, we did perform post hoc recognition tests for the Probe faces, to determine how well each Probe face was seen. The results are summarised in figure 5.1 of Hajilou [24], where qualitatively ratings of how often a Probe face was seen reduced across blocks, which is consistent with them becoming less salient. However, a one-way anova with factor block of three levels was not significant. For further details, see Hajilou [24].

4.8 Interpretation

The findings in this paper can be related to previously identified electrophysiological features and existing theories. To help making this connection, scalp maps and stack plots for our ERP effects can be found in Appendix 3.

Target P3b: the Target elicited a clear P3b pattern (see, for example, Appendix 3, Fig. 12). This begins around Fz and moves posteriorly as time progresses (see right side of Fig. 12). As frequently observed, this component peaks later in RSVP than in standard Odd-Ball paradigms, here reaching its maximum around 550 ms post stimulus onset in the grand average, considerably later than the 300 ms associated with the P300 name also used for the Odd-ball P3b component.

This is a pattern we have observed consistently for the Target across our RSVP experiments, when using a mastoid reference (e.g. [1, 2, 12, 13, 24, 25]). This fits well with the interpretation that task relevance of the eliciting stimulus ensures the P3b amplitude is large. This P3b pattern for the Target also resonates with theories of conscious access, which view the P3b as the electrophysiological correlate of conscious perception (e.g. [20, 40]), what we have called the Breakthrough-P3 (e.g. [1]). This said, consistent with what we have just discussed, the P3b seems to be modulated by other cognitive characteristics beyond conscious access, such as task relevance.

Probe: One interpretation that we have for the electrical response we observe for the Probe is that we are again seeing what we have called a drunken-wavelet pattern (see [1]), possibly sat on top of a P3. The time–frequency results we observe here and in [1] are very similar, being over a similar frequency range, largely focussed on theta, and over a similar time-course. Additionally, the scalp maps through time observed for the famous-faces drunken-wavelet (see [24], figure 4.5) also show a strikingly similar spatial distribution to that which we have observed here, see Fig. 11 in Appendix 3. This said, the oscillatory feature in this paper is considerably weaker than in [1], which is presumably because familiar faces are not as salient as famous faces. Interestingly, such a (stimulus-locked) theta burst oscillation has also been observed with similar timing for the Odd-ball P3 [4], suggesting that such a burst could be a more general feature of the P3-complex.

Linking this Probe-elicited theta feature to face detection findings has some uncertainty, since the Target did not elicit such a feature, or if it is present, it is much weaker. However, the Target is also a face.

A striking observation of the findings here, which again was also present in [1], is that if a P3b is present, it is weak. This somewhat clashes with the conscious access interpretation of the P3b, since the Probe is certainly consciously perceived (although probably not as frequently and perhaps more fleetingly than the Target). This said, it is possible that the P3b is present for the Probe, but it is obscured by the presence of the theta burst, with the possibility that the maximum point of the P3b coincides with a negativity of the theta burst. Further work could explore these questions by using a larger electrode montage than employed in this paper, applying source separation techniques (such as ICA, PCA and Empirical Mode Decomposition in the frequency domain) and undertaking source localisation to identify the brain sources of different features of the scalp recorded EEG.

4.9 Theories of familiar face components

Our time-domain analyses focus on the N400f and P600f face components, which are reported in the literature as correlates of the brain's processing of familiar faces [6, 22, 23, 51]. A question is how consistent the components we observe are to those reported in the literature, especially since the classic experiments that have characterised the face N400f and P600f are quite different to ours, e.g. they are not RSVP experiments.

We can though only inform this question, since, as previously discussed, our experimental designs were not set-up for a high-density characterisation of electrical fields on the scalp, although, a companion study to that reported here did collect 32 electrodes (see [24], chapter 7) and we will refer to those findings. We focus on results for the Probe, since this is the stimulus that was pre-experimentally familiar.

Considering the classic N400f and P600f literature, figure 5 and 6 of Eimer [22], figure 2 of Eimer [23] and figure 2 of Bentin & Deouell [6], the N400f and P600f are broadly distributed across the scalp, and typically relatively symmetric around the centreline, with N400f probably biggest at Cz and P600f between Cz and Pz. Our N400f and P600f effects also look to be symmetric on the scalp, and typically largest on the centreline. Our topography is perhaps slightly more anterior (see scalp maps in Fig. 11 and Figure 7.4 in [24], chapter 7). That is, our effects are more Fz-Cz than the Cz-Pz distribution classically reported for N400f and P600f, although our statistical effects were not bigger at Fz—see Table 7 in the appendix.

Thus, the spatial profiles of our N400f and P600fs do have similarities to the classical characterisation of these effects. However, it is notable that employing a time–frequency analysis identifies a feature that is consistent in terms of frequency range across an extended period of time (e.g. see Figs. 6 and 7, leftmost two plots). This seems to suggest a multicycle (stimulus-locked) burst at a consistent frequency (theta), rather than a sequence of separate evoked components. (This is what we called the drunken wavelet in Alsufyani et al. [1], and discussed in the previous subsection).

In fact, we have observed this extended oscillatory feature in a number of experiments, including in unpublished data for a letters-in-digits RSVP experiment, i.e. even when faces are not presented, and, as previously discussed, the feature has even been observed with auditory stimuli in an odd-ball P3 task [4]. Thus, it may be that this oscillatory feature is more general than just faces or even RSVP experiments.

4.10 Broader context

Evidence for the effectiveness of the Fringe-P3 method as a concealed information test (i.e. the RSVP-P3 CIT) is now accumulating across a range of studies, both from Bowman's lab (e.g. [1, 2, 12, 13, 24, 25]) and other labs, (e.g. [16, 42, 43, 49, 50]). These studies have employed a variety of physiological measurement techniques and analysis methods, e.g. pupillometry [16, 42, 43], neural network classification [49], and multimodal classification [50].

The work by Chen et al. and Rosenzweig & Bonneh seeks to identify a change in a feature observable at the eye that correlates with occurrence of the P3-complex. Potential features are the rate of micro-saccades, rate of eye-blinks or pupil size. This work has the potential to greatly increase the practical applicability of the Fringe-P3 method and in particular, the RSVP-CIT. Specifically, the work raises the possibility that the method could be applied with a simple camera, without requiring electrodes to be placed on the scalp. However, while strong group level effects have been observed, although present, effects at the per-individual level have in general been weaker than those observed with EEG. Work to try to increase the effectiveness of these eye-related measures is an important direction for future research.

The work by Wang, Qi et al. and Wang, van der Mijn et al. advanced the classification procedures applied to Fringe-P3 EEG data. Wang, Qi et al. used neural networks and Wang, van der Mijn et al. combined EEG and pupillometry measures in a multi-modal inference procedure. These approaches all speak to the potential that more refined EEG analysis methods offer to increase classification accuracy; although typically, the effectiveness of modern AI methods, which are proving so successful in other contexts, is dependent upon data set size [44], which is typically small with neuroimaging, such as EEG.

4.11 The fringe-P3 method

It is notable that the form of the evoked response that we observe for probe stimuli does vary according to the type of stimulus presented. In particular, for face stimuli, we observe an oscillatory pattern in the theta frequency range. This, for example, is evident in the green time series in Fig. 2, and is an even more pronounced feature in our famous faces experiment [1]. The practical value of the Fringe-P3 method is not impacted by the presence or absence of an actual P3 in these evoked response patterns; indeed, as we have shown here and in [1], analyses that use this oscillatory feature are particularly effective. However, these heterogeneous evoked patterns might challenge the reference to P3 in the name we have attributed to this method (i.e. Fringe-P3). However, for the Probe (green in Fig. 2), if one filtered out high frequencies (which amongst other things will make the negative deflection just after 600 ms less sharp), and then subtracted out a (negative going) wavelet, centred on the negative deflection just before 400 ms (which our famous faces data [1] suggests is the central phenomenon of the wavelet), one might end up with something that looks like a (weak) P3, peaking around the same time as the Target (red in Fig. 2). Thus, although further work is certainly required, it may be that the oscillatory feature in our faces experiment is "sat on top of" a P3. This would justify the term Fringe-P3 method still being applicable.

This paper's experiment extended our earlier work, which, together, demonstrated that both highly evocative faces (i.e. famous faces experiment) and personally familiar faces (i.e. this experiment's lecturers) can be employed in RSVP-based Fringe-P3 studies, and that highly familiar faces can breakthrough into conscious awareness, on a per-individual basis. (Note, his link between the P3 and conscious perception has been substantiated in Pincham et al. [40]). The results raise the possibility that one could apply our findings to the differentiation of deceivers and non-deceivers, in the application of crime compatriots, whereby, a suspect's familiarity with a criminal/terrorist can be established using faces.

	Moments of pixel intensity distribution, etc						Central spatial moments, etc											
		Standard	Pixel intensities					Centre of Mass		Variance			Skewness		Kurtosis			
Image	**Mean**	**Deviation**	**Min**	**Max**	**Skewness**	**Kurtosis**	**Mass**	**X**	**Y**	**X**	**Y**	**X/Y**	**X**	**Y**	**X**	**Y**	**Orient.**	**Elong.**
1	177.152	70.06	16	254	-0.855	-0.92	54420984	314.149	226.752	39893.559	17393.776	-1279.867	0.04	-0.002	-1.393	-1.111	-3.245	2.79E+16
2	187.353	67.91	5	236	-1.327	0.425	57554800	320.169	215.014	39347.989	17162.463	82.56	0.002	0.115	-1.382	-1.142	0.213	2.83E+16
3	193.767	56.193	19	234	-1.185	0.011	59525097	319.386	232.602	39580.758	19290.221	-279.766	0.005	0.038	-1.417	-1.205	-0.79	2.45E+16
4	198.618	52.519	14	247	-1.416	0.61	61015386	323.088	229.378	38840.365	19410.628	237.678	-0.041	0.086	-1.39	-1.21	0.701	2.31E+16
5	187.79	60.742	19	232	-1.082	-0.177	57689191	323.132	232.778	39948.209	18421.742	199.96	-0.041	-0.004	-1.411	-1.175	0.532	2.67E+16
6	195.817	50.41	34	232	-0.888	-0.896	60155100	321.927	230.678	39055.425	19225.214	217.597	-0.025	0.047	-1.394	-1.214	0.629	2.37E+16
7	192.487	52.191	22	234	-0.782	-1.05	59132112	320.787	232.453	39606.576	19022.087	198.644	-0.007	0.025	-1.414	-1.194	0.553	2.51E+16
8	186.771	64.966	14	232	-1.098	-0.217	57376200	318.197	228.039	40012.504	18274.831	-266.609	0.013	0.013	-1.417	-1.165	-0.703	2.71E+16
9	187.751	59.178	19	235	-1.084	-0.025	57677256	316.362	227.192	38659.642	18441.122	-544.455	0.033	0.035	-1.36	-1.176	-1.541	2.37E+16
10	189.362	58.672	24	251	-1.016	-0.494	58171905	323.762	239.93	40295.691	19872.835	62.168	-0.051	-0.04	-1.434	-1.227	0.174	2.43E+16
11	190.695	58.007	25	250	-1.019	-0.504	58581446	320.102	226.514	39174.026	18383.798	273.989	-0.002	0.071	-1.386	-1.155	0.755	2.53E+16
12	187.501	59.047	18	232	-0.893	-0.782	57600214	321.267	224.287	39874.718	18504.654	226.703	-0.019	0.07	-1.416	-1.197	0.608	2.63E+16
	189.589	**59.158**	**19.083**	**239.083**	**-1.054**	**-0.335**	**58241641**	**320.194**	**228.801**	**39524.122**	**18616.948**	**-72.617**	**-0.008**	**0.038**	**-1.401**	**-1.181**	**-0.176**	**2.55E+16**
A	190.677	57.078	20	234	-0.952	-0.594	58576046	315.439	223.58	39216.753	18579.933	-496.141	0.052	0.095	-1.388	-1.184	-1.376	2.50E+16
B	189.188	55.626	26	242	-0.856	-0.759	58118457	314.352	227.289	39708.714	19714.327	-564.519	0.07	0.119	-1.397	-1.207	-1.616	2.33E+16
C	190.958	58.428	23	233	-1.06	-0.281	58662159	316.972	230.788	39381.668	18995.544	-465.001	0.033	0.011	-1.394	-1.207	-1.306	2.44E+16
D	197.537	53.197	24	250	-1.221	-0.056	60683282	320.306	229.788	38996.207	19278.057	-26.596	-0.006	0.063	-1.397	-1.215	-0.077	2.36E+16
E	185.631	57.873	32	234	-0.72	-1.105	57025898	315.628	228.49	40292.116	19131.821	-466.279	0.05	0.061	-1.422	-1.181	-1.262	2.56E+16
F	198.466	52.552	36	234	-1.417	0.806	60968809	317.997	232.829	38365.585	18589.158	129.116	0.026	-0.009	-1.362	-1.187	0.374	2.39E+16
G	171.577	70.262	17	232	-0.592	-1.248	52708561	321.033	226.99	42299.482	19365.091	577.51	-0.012	0.075	-1.481	-1.191	1.442	2.78E+16
Average	**189.148**	**57.859**	**25.429**	**237.000**	**-0.974**	**-0.462**	**58106173**	**317.390**	**228.536**	**39751.504**	**19093.419**	**-187.416**	**0.030**	**0.059**	**-1.406**	**-1.196**	**-0.546**	**2.48E+16**
p-value	***0.909***	***0.656***	***0.071***	***0.564***	***0.526***	***0.686***	***0.909***	***0.044***	***0.899***	***0.664***	***0.106***	***0.593***	***0.020***	***0.329***	***0.768***	***0.188***	***0.519***	***0.377***
t(18)	**0.134**	**0.452**	**-1.927**	**0.548**	**-0.735**	**0.448**	**0.134**	**2.144**	**0.109**	**-0.556**	**-1.453**	**0.536**	**-2.727**	**-1.027**	**0.356**	**1.124**	**0.653**	**0.880**
ind. t-test uneq																		

key: *12 x Probes* / *7 x Irrelevants*

Central Spatial Moments, etc
Mass = sum of all intensities
Center of Mass = location of centre of mass in x and y (1st order spatial moment)
Variance = Variance (spread) around centre of mass in x, y and xy (2nd order spatial moment)
Skewness = Coefficient of skewness (symmetry) in x and y axis (3rd order spatial moments)
Kurtosis = Kurtosis (flatness) (4th order spatial moments)
Orient. = For elongated objects, describes the orientation (in degrees) of the "long" direction with respect to horizontal (x axis)
Elong. = Elongation (Eccentricity) is the ratio of longest to shortest distance vectors from the object's centroid to its boundaries

Fig. 10 Low level visual features: twelve Probe faces and seven Irrelevant faces are analysed with moments of pixel intensity distribution and central spatial moments. P-values of statistical tests (two sample two-tailed independent t-tests) comparing Probe and Irrelevant faces are shown in the yellow highlighted row

Table 6 Two-tailed two-sample independent t-tests comparing Probe and Irrelevant faces across 11 summarised spatial frequency bins, increasing from low to high spatial frequencies

Spatial frequency bands											
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
t(18)	0.737209	0.013822	−0.26863	−0.43644	−0.447	−0.29927	−0.10674	0.022069	0.160723	0.425134	0.522701
p-value	0.470497	0.989124	0.791266	0.667706	0.66021	0.768164	0.916179	0.982635	0.874102	0.675778	0.607557

No p-values are close to showing a reliable statistical difference

Table 7 All group-level effects in article summarised

Group-level effects, difference between probe and irrelevant	*p*-value
at Pz	
Time domain P600f: t(13) = 3.3198, p = 0.0055, d = 1.3691.	0.0055
Time domain N400f: t(13) = − 1.9035, p = 0.0794, d = 0.4532.	0.0794
Full frequency-range (0.5 to 45 Hz) for ERSP t(13) = 2.9737, p = 0.0108, d = 1.0417	0.0108
Full frequency-range (0.5 to 45 Hz) for ITC t(13) = 1.8064, p = 0.0941, d = 0.7997	0.0941
Narrower frequency-band (0.5 to 7 Hz) for ERSP t(13) = 3.4769, p = 0.0041, d = 1.2649.	0.0041
Narrower frequency range (0.5 to 7 Hz), for ITC: t(13) = 2.322, p = 0.0371, d = 0.7442	0.0371
Fz	
Time domain P600f: t(13) = 3.3151, p = 0.0056, d = 0.9834	0.0056
Cz	
Time domain P600f, t(13) = 2.8346, p = 0.0141, d = 0.96494	0.0141

FDR not significant:	(dark grey)
FDR significant:	(light shading)

Table 8 Results of Benjamini–Hochberg procedure applied to results shown in Table 7

False discovery rate					
FDR = 0.05					
p-values	Sorted	Rank	No. tests	BH critical val	Decision
0.0055	0.0041	1	8	0.00625	Candidate
0.0794	0.0055	2		0.0125	Candidate
0.0108	0.0056	3		0.01875	Candidate
0.0941	0.0108	4		0.025	Candidate
0.0041	0.0141	5		0.03125	Candidate
0.0371	0.0371	6		0.0375	Candidate
0.0056	0.0794	7		0.04375	Not significant
0.0141	0.0941	8		0.05	Not significant

FDR not significant: (darker grey)

FDR significant: (light grey)

First column are p-values of the article in order of experiments, as per last column of Table 7. Second column shows column one p-values sorted from small to large, with result of FDR shown in grey (light and darker). Fifth column shows the critical value calculated using the Benjamini–Hochberg method associated with each rank position. Last column indicates which p-value is below its corresponding critical value

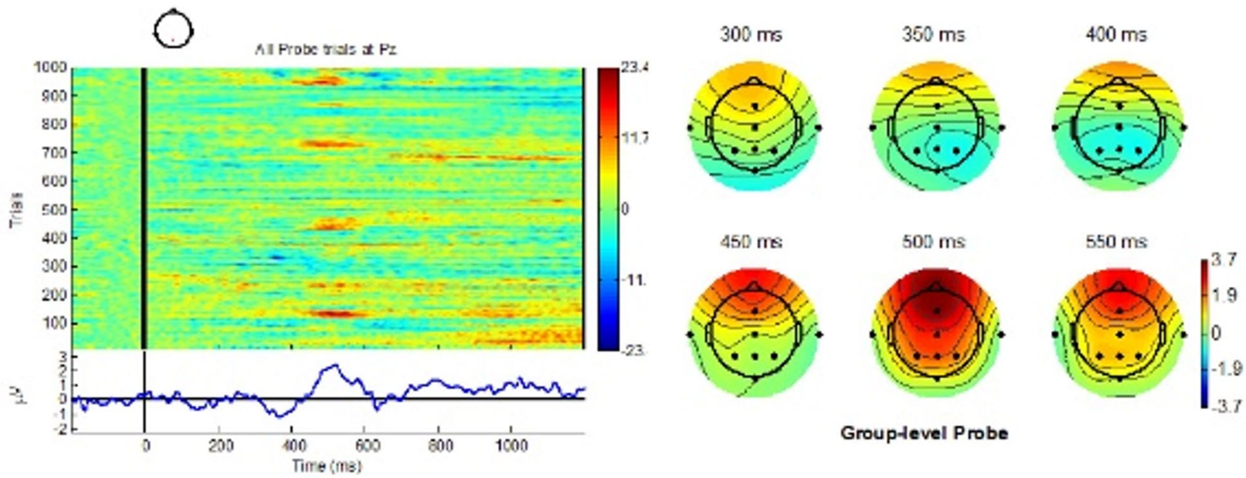

Fig. 11 Group-level view of all (1007) Probe trials, over time, at Pz (left plot, grand average below) and interpolated scalp topography of the Probe (right plot, which must be treated with caution, due to the small number of electrodes). An oscillatory pattern can be observed, with lowest negativity observed at 350 ms (a weak N400f), and highest positivity at 500 ms. The scalp map scale ranges from −3.7 to + 3.7 µV, which is lower than the scale for Target (see Fig. 12)

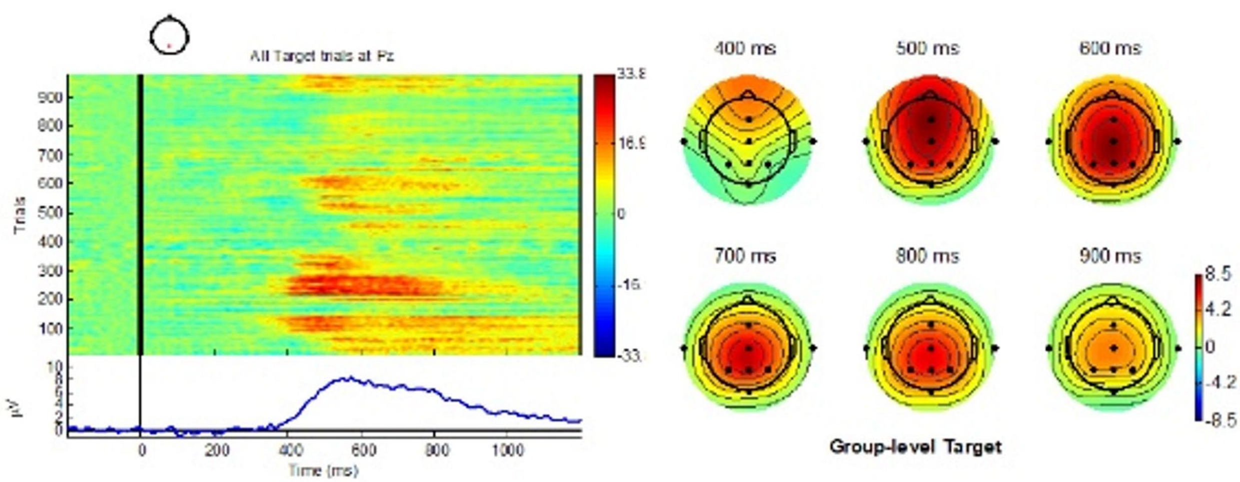

Fig. 12 Group-level view of all (987) Target trials, in order of appearance over time, at Pz (left plot, grand average below), and the corresponding scalp topography of the Target (right plot), showing a prevailing positivity, peaking at around 550 ms, with the electrical field moving posteriorly through time. Having used a limited number of (8) electrodes in this experiment, it must be noted that MATLAB employs an interpolation algorithm to represent the full scalp pattern. Therefore, estimated electric potential values are used at scalp locations between the actual recording sites, and the presented scalp topographies carry considerable uncertainty, especially in respect of laterality of effects, since we have few electrodes beyond the central line. The scalp map scale ranges from −8.5 to + 8.5 µV

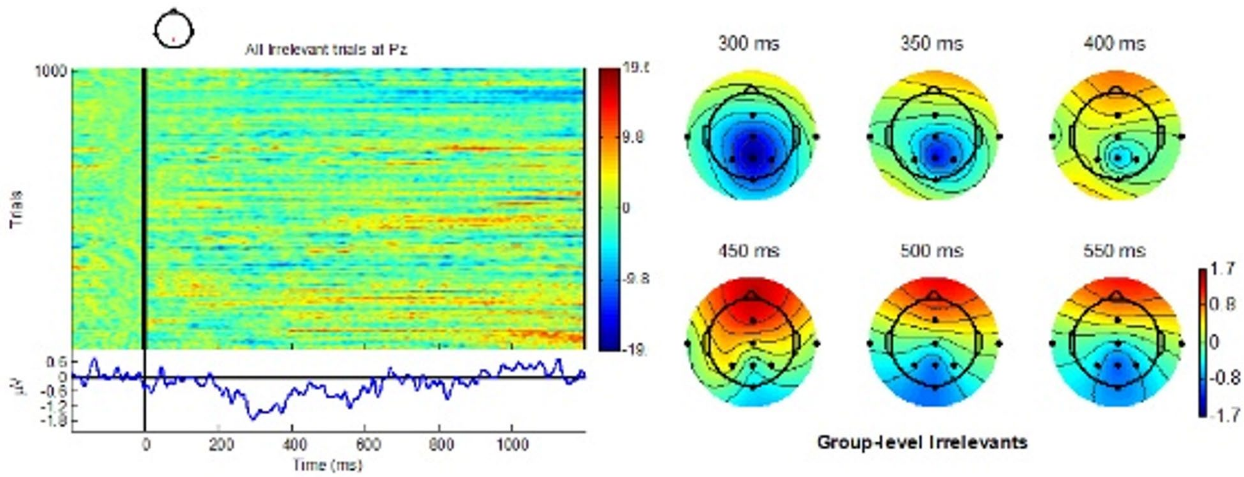

Fig. 13 Group-level view of all (1017) Irrelevant trials, over time at Pz (left side, grand average below), and the interpolated scalp map representation of the Irrelevant (right plot, which must be treated with caution, due to the small number of electrodes), showing a distinctly different pattern when compared to the Probe (albeit, the small negativity, peaking at 300 ms, was not expected). The scalp map scale ranges from − 1.7 to + 1.7 μV, which is lower than the scale for Probe (see Fig. 11)

Appendix 1: Low-level visual features of faces

An important potential confound to rule out is that our Probe and Irrelevant faces have different low-level visual features. If they did, these features could explain why Probes breakthrough into awareness while Irrelevants do not. We assessed this using three classes of image statistic:

1. *Moments of pixel intensity distribution*: these are the four moments of the probability distribution constructed from all the pixel intensities in an image. These are key summary statistics of these distributions. In particular, the first moment is the mean luminance and the second moment (the variance) is a good proxy for the contrast of an image.
2. *Central moments of image*: these are *spatial* moments calculated relative to the centre of mass of an image. They are weighted averages of intensity relative to the centre of mass of the image. Central moments can act as summary measures of spatial frequency. Low-order central moments reflect low frequencies (smooth features), while higher-order moments reflect high frequencies (sharp boundaries and fine detail).
3. *Summaries of spatial frequencies*: to be absolutely certain there were not differences in spatial frequencies, we also tested differences between Probe and Irrelevant faces for direct summary measures of spatial frequencies.

The results of the first two of these analyses are shown in Fig. 10. We have measured 18 quantities across Probe and Irrelevant faces and run 18 statistical tests, with a false positive rate of one in 20. We obtained two significant effects: a just significant effect on the X dimension of Centre of Mass ($p = 0.044$) and a stronger effect on the X dimension of Central Spatial Moment Skewness ($p = 0.02$). Neither of these are measures that could be considered plausible reasons why Probe and Irrelevant faces would be perceived so differently.

Spatial Frequency: As already discussed, we also investigated whether there were differences in spatial frequency bands between Probe and Irrelevant faces. Specifically, we generated two-dimensional spatial frequency spectra for each Probe and Irrelevant face.

Then, we summarised each such 2d-spectra into a one-dimension structure, by azimuthal (radial) averaging. This left us with 385 averages along this one-dimension, from lowest to highest spatial frequencies. We then binned (down-sampled) this one-dimension, by averaging in contiguous bins of 35 entries, leaving us with 11 values from low to high frequency bins for each face.

We then ran two-tailed two-sample independent t-tests to look for a statistical difference between Probe and Irrelevant in each of the 11 bins. The results of these tests are shown in Table 6. None of these 11 tests approached the alpha level, giving us no evidence of a reliable difference between Probe and Irrelevant faces. (The clustering of p-values at high values likely reflects that the 11 bins are not statistically independent of each other, since a spatial frequency spectrum exhibits smoothness across frequencies.)

We believe that these findings support the central argument of this paper that the Probes are consciously perceived, generating associated evoked responses, and Irrelevants are not, because Probes are familiar and Irrelevants are not.

Appendix 2: Multiple comparisons

This article contains several group-level effects. In this appendix, we asses our vulnerability to type I errors, by performing a multiple comparisons correction across all the group-level effects reported. We ran a false discovery rate correction based upon the Benjamini–Hochberg method [5]. Table 7 summarises the group-level effects of this article. Table 8 presents the false discovery rate analysis, with results that survive the correction shown in light grey. These results are also indicated in the final column of Table 7. In fact, because of the strength of our group-level effects, we do not lose any significant effects.

Appendix 3: Stack plots and scalp maps

We present further information on the ERP effects we consider in the main body. We show stack plots (also called ERP images) of single trials (with grand average below) on the left side of our figures and electrical fields through time on the right side: Fig. 12 depicts the Target condition; Fig. 11 the Probe; and Fig. 13 the Irrelevant. In particular, Fig. 12 shows a classic P3b pattern. Figure 11 shows that the Probe pattern (for an early time period, in which the effect is biggest) is considerably more frontal than the classic P3b pattern (e.g. scalp maps at 500 ms and 550 ms). Although, inferences on scalp maps should be made with caution, since we only have eight electrodes. This is particularly with regard to the laterality of the effect.

Despite these topographic patterns, we have maintained our focus on the Pz electrode, since this was a priori the most expected and we do not wish to inflate false positive rates (i.e. Type I error rates) by tailoring our analysis post hoc to the observed data. This said, the findings presented here should serve as prior precedents for analyses at the Fz electrode with these classes of stimuli in future studies.

Acknowledgements
N/A.

Author contributions
H.B. directed the research. H.B. and O.H. designed the experiment. O.H. coded the experiment and collected the data. H.B. and O.H. wrote the main manuscript text and reviewed the manuscript.

Funding
This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability
The data for this experiment will be made available on a neuroscience repository.

Declarations

Ethics approval and consent to participate
The Sciences Research Ethics Advisory Group, at the University of Kent, approved the study, under approval 0671314. All participants provided informed consent for the study and use of their data.

Competing interests
The authors declare no competing interests.

Received: 22 July 2025 / Accepted: 11 April 2026
Published online: 14 May 2026

References

1. Alsufyani A, Hajilou O, Zoumpoulaki A, Filetti M, Solomon J, Bowman HC (2019) Breakthrough percepts of famous faces. Psychophysiology 56(1):e13279. https://doi.org/10.1111/psyp.13279
2. Alsufyani A, Harris K, Zoumpoulaki A, Filetti M, Bowman H (2021) Breakthrough percepts of famous names. Cortex 139:267–281
3. Avilés A, Bowman H, Wyble B (2020) On the limits of evidence accumulation of the preconscious percept. Cognition 195:104080
4. Başar-Eroglu C, Demiralp T (2001) Event-related theta oscillations: an integrative and comparative approach in the human and animal brain. Int J Psychophysiol 39(2–3):167–195
5. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Roy Stat Soc: Ser B (Methodol) 57(1):289–300
6. Bentin S, Deouell LY (2000) Structural encoding and identification in face processing: ERP evidence for separate mechanisms. Cognitive Neuropsychol 17:35–54
7. Bowman H, Assant A, Aviles A (2025) The Missing Memory Imprint. bioRxiv, 2025–05.
8. Bowman H, Avilés A (2021) Fragile Memories for Fleeting Percepts. PsyArXiv, https://doi.org/10.31234/osf.io/fr3bg.
9. Bowman H, Avilés A (2022) No subliminal memory for spaced repeated images in rapid-serial-visual-presentation streams of thousands of images. Psychol Sci 33(11):1959–1965
10. Bowman H, Wyble B (2007) The simultaneous type, serial token model of temporal attention and working memory. Psychol Rev 114(1):38
11. Bowman H, Brooks JL, Hajilou O, Zoumpoulaki A, Litvak V (2020) Breaking the circularity in circular analyses: simulations and formal treatment of the flattened average approach. PLoS Comput Biol 16(11):e1008286
12 Bowman H, Filetti M, Alsufyani A, Janssen D, Su L (2014) Countering countermeasures: detecting identity lies by detecting conscious breakthrough. PLoS ONE 9(3):e90595
13. Bowman H, Filetti M, Janssen D, Su L, Alsufyani A, Wyble B (2013) Subliminal salience search illustrated: EEG identity and deception detection on the fringe of awareness. PLoS ONE. https://doi.org/10.1371/journal.pone.0054258
14. Bowman H, Wyble B, Chennu S, Craston P (2008) A reciprocal relationship between bottom-up trace strength and the attentional blink bottleneck: relating the LC–NE and ST2 models. Brain Res 1202:25–42
15. Brooks J, Zoumpoulaki A, Bowman H (2017) Data-driven region-of-interest selection without inflating Type I error rate: safe data-driven ROI selection. Psychophysiology 54:100–113. https://doi.org/10.1111/psyp.12682
16. Chen IY, Karabay A, Mathôt S, Bowman H, Akyürek EG (2023) Concealed identity information detection with pupillometry in rapid serial visual presentation. Psychophysiology 60(1):e14155
17. Chennu S, Alsufyani A, Filetti M, Owen AM, Bowman H (2013) The cost of space independence in P300-BCI spellers. J Neuroeng Rehabil 10(1):1–13
18. Craston P, Wyble B, Chennu S, Bowman H (2009) The attentional blink reveals serial working memory encoding: evidence from virtual and human event-related potentials. J Cogn Neurosci 21(3):550–566
19 Curan T, Hancock J (2007) The FN400 indexes familiarity-based recognition of faces. Neuroimage 36(2):464–471
20 Dehaene S, Changeux JP, Naccache L (2011) The global neuronal workspace model of conscious access: from neuronal architectures to clinical applications. In: Dehaene S, Christen Y (eds) Characterizing consciousness: from cognition to the clinic? Springer, Berlin, Heidelberg, pp 55–84
21. Delorme A, Makeig S (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component

analysis. J Neurosci Methods 134(1):9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009
22. Eimer M (2000) Effects of face inversion on the structural encoding and recognition of faces: evidence from event-related brain potentials. Cogn Brain Res 10(1–2):145–158
23. Eimer M (2000) Event-related brain potentials distinguish processing stages involved in face perception and recognition. Clin Neurophysiol 111:694–705. https://doi.org/10.1016/S1388-2457(99)00285-0
24. Hajilou O (2020) Face Perception and Recognition, on the Fringe of Human Awareness. Doctor of Philosophy (PhD) thesis, University of Kent. https://doi.org/10.22024/UniKent/01.02.85577
25. Harris K, Miller C, Jose B, Beech A, Woodhams J, Bowman H (2021) Breakthrough percepts of online identity: detecting recognition of email addresses on the fringe of awareness. Eur J Neurosci 53(3):895–901
27. Hellerstedt R, Moccia A, Brunskill CM, Bowman H, Bergström ZM (2021) Aging reduces EEG markers of recognition despite intact performance: implications for forensic memory detection. Cortex 140:80–97
28. Jasper H (1958) The ten-twenty electrode system of the International Federation. Electroencephalogr Clin Neurophysiol 10:371–375
29. Johnson M, Rosenfeld J (1992) Oddball-evoked P300-based method of deception detection in the laboratory II: utilization of non-selective activation of relevant knowledge. Int J Psychophysiol 12(3):289–306
30. Kaufmann T, Schulz S, Grünzinger C, Kübler A (2011) Flashing characters with famous faces improves ERP-based brain-computer interface performance. J Neural Eng 8:056016. https://doi.org/10.1088/1741-2560/8/5/056016
31. Kilner J (2013) Bias in a common EEG and MEG statistical analysis and how to avoid it. Clin Neurophysiol 124(10):2062–2063
32. Kriegeskorte N, Simmons W, Bellgowan P, Baker C (2009) Circular analysis in systems neuroscience: the dangers of double dipping. Nat Neurosci 12(5):535–540. https://doi.org/10.1038/nn.2303
33. Labkovsky E, Rosenfeld J (2012) The P300-based, complex trial protocol for concealed information detection resists any number of sequential countermeasures against up to five irrelevant stimuli. Appl Psychophysiol Biofeedback 37:1. https://doi.org/10.1007/s10484-011-9171-0
34. Lefebvre C, Marchand Y, Smith S, Connolly J (2007) Determining eyewitness identification accuracy using event-related brain potentials (ERPs). Psychophysiology 44(6):894–904
35 Lukács G, Weiss B, Dalos V, Kilencz T, Tudja S, Csifcsák G (2016) The first independent study on the complex trial protocol version of the P300-based concealed information test: corroboration of previous findings and highlights on vulnerabilities. Int J Psychophysiol 110:56–65
36. Makeig S, Debener S, Onton J, Delorme A (2004) Mining event-related brain dynamics. Trends Cogn Sci 8(5):204–210
37 Meijer E, Smulders F, Wolf A (2009) The contribution of mere recognition to the P300 effect in a concealed information test. Appl Psychophysiol Biofeedback 34:221
38. Meixner J, Haynes A, Winograd M, Brown J, Rosenfeld J (2009) Assigned versus random, countermeasure-like responses in the p300 based complex trial protocol for detection of deception: task demand effects. Appl Psychophysiol Biofeedback 34(3):209–220
39. Minear M, Park D (2004) A lifespan database of adult facial stimuli. Behav Res Method Instrum Comput J Psychonomic Soc Inc 36:630–633
40. Pincham HL, Bowman H, Szucs D (2016) The experiential blink: mapping the cost of working memory encoding onto conscious perception in the attentional blink. Cortex 81:35–49
41. Rosenfeld P, Soskins M, Bosch G, Ryan A (2004) Simple, effective countermeasures to P300-based tests of detection of concealed information. Psychophysiology 41(2):205–219
42. Rosenzweig G, Bonneh YS (2020) Concealed information revealed by involuntary eye movements on the fringe of awareness in a mock terror experiment. Sci Rep 10(1):14355
43. Rosenzweig G, Bonneh YS (2019) Familiarity revealed by involuntary eye movements on the fringe of awareness. Sci Rep 9(1):3029
44. Saranti M, Neville D, White A, Rotshtein P, Hope TM, Price CJ, Bowman H (2025) Predicting language outcome after stroke using machine learning: in search of the big data benefit. Neuroimage Clin. https://doi.org/10.1016/j.nicl.2025.103858
45. Sutton S, Braren M, Zubin J, John E (1965) Evoked-potential correlates of stimulus uncertainty. Science 150(3700):1187–1188
46. Touryan J, Gibson L, Horne JH, Weber P (2011) Real-time measurement of face recognition in rapid serial visual presentation. Front Psychol 2:42
47. Trenner MS, Jentzsch I, Sommer W (2004) Face repetition effects in direct and indirect tasks: an event-related brain potentials study. Brain Res Cogn Brain Res 21:388–400
48. Wang Y, Jung T (2010) Visual stimulus design for high-rate SSVEP BCI. Electron Lett 46:1057-U27
49. Wang H, Qi Y, Yu H, Wang Y, Liu C, Hu G, Pan G (2021) RCIT: an RSVP-based concealed information test framework using EEG signals. IEEE Trans Cogn Dev Syst 14(2):541–551
50. Wang S, van der Mijn R, Karabay A, Akyurek E (2025) Multimodal detection of concealed information in rapid serial visual presentation using EEG and pupillometry. OSF, PsyArXiv, https://osf.io/preprints/psyarxiv/jaekf_v1.
51. Wiese H, Schweinberger SR, Kovács G (2024) The neural dynamics of familiar face recognition. Neurosci Biobehav Rev 167:105943

Publisher's Note