Breakthrough percepts of famous names

Alsufyani, Abdulmajeed; Harris, Kathryn; Zoumpoulaki, Alexia; Filetti, Marco; Bowman, Howard

DOI:
10.1016/j.cortex.2021.02.030

Breakthrough Percepts of Famous Names

Abdulmajeed Alsufyani*[1], Kathryn Harris*[2], Alexia Zoumpoulaki[3], Marco Filetti[4], and Howard Bowman[2, 5]

*These authors contributed equally to this work.

[1] Department of Computer Science, Taif University, Taif, Kingdom of Saudi Arabia.

[2] School of Computing, University of Kent, Canterbury, United Kingdom

[3] School of Computer Science and Informatics, Cardiff University, Cardiff, United Kingdom.

[4] Helsinki Institute for Information Technology, University of Helsinki, Helsinki, Finland.

[5] School of Psychology, University of Birmingham, Birmingham, United Kingdom

CRediT Author Statement:

Abdulmajeed Alsufyani: Conceptualisation, Methodology, Software, Investigation, Writing

Kathryn Harris: Software, Formal Analysis, Data Curation, Writing, Visualization

Alexia Zoumpoulaki: Methodology

Marco Filetti: Software, Methodology

Howard Bowman: Conceptualisation, Methodology, Supervision, Project Administration, Writing

Corresponding Author:

Kathryn Harris, School of Computing, University of Kent, Canterbury, Kent, CT2 7NF, United Kingdom

Email: kh468@kent.ac.uk

Abstract

Studies have shown that presenting own-name stimuli on the fringe of awareness in Rapid Serial Visual Presentation (RSVP) generates a P3 component and provides an accurate and countermeasure resistant method for detecting identity deception (Bowman et al., 2013, 2014). The current study investigates how effective this Fringe-P3 method is at detecting recognition of familiar name stimuli with lower salience (i.e. famous names) than own-name stimuli, as well as its accuracy with multi-item stimuli (i.e. first and second name pairs presented sequentially). The results demonstrated a highly significant ERP difference between famous and non-famous names at the group level and a detectable P3 for famous names for 86% of participants at the individual level. This demonstrates that the Fringe-P3 method can be used for detecting name stimuli other than own-names and for multi-item stimuli, thus further supporting the method's potential usefulness in forensic applications such as in detecting recognition of accomplices.

Keywords: EEG, Rapid Serial Visual Presentation, Concealed Knowledge Test, Fringe-P3, Famous Names

1. Introduction

Since Lawrence in 1971, Rapid Serial Visual Presentation (RSVP) has been extensively used in the scientific study of perception and attention (Bowman & Wyble, 2007; Broadbent & Broadbent, 1986; Raymond et al., 1992). However, its practical use in cognitive systems applications has only recently come to the fore. This is especially true of the combination of RSVP with EEG, which offers a very high presentation bandwidth along with a means to detect what a participant found salient. This combination involves stimuli presented very rapidly, one after the other, at a fixed location, with EEG used to determine which of these stimuli "breaks through" into awareness. Specifically, stimuli that are perceived and encoded into working memory induce a P3 component (Craston et al., 2009; Martens et al., 2006; Vogel et al., 1998) and this perception is preferential, in that the stimuli perceived are typically the ones that are salient to the perceiving brain. We have described the cognitive process engaged by RSVP as Subliminal Salience Search (SSS) (Bowman et al., 2013). This is in the sense that the brain is searching amongst the presented stimuli for those that are of interest to the brain, i.e. are salient, and that the search is subliminal, since very few of the stimuli are sufficiently perceived to be encoded into working memory (Bowman & Avilés, in preparation). Thus, the method can be used to determine a participant's preferences from their electrical brain activity. Furthermore, these preferences could reflect an explicit task set or an implicit (intrinsic) salience. For example, the RSVP P3-speller (Acqualagna & Blankertz, 2013; Chennu et al., 2013) involves the former of these: participants search for a specific letter under the guidance of cognitive control and their brain detects its presence because it is relevant to the current task, not because of any fundamental intrinsic salience. In contrast, the Fringe-P3 method for detecting concealed knowledge (Alsufyani et al., 2019; Bowman et al., 2013, 2014) relies upon the stimuli (the concealed knowledge, e.g. the participant's own-name) being intrinsically salient, causing it to be selected by the brain and consciously perceived.

Indeed, the Fringe-P3 method was shown in Bowman et al. (2013) to be an effective means to detect identity deception. This is because in the RSVP streams of names presented, a participant's own-name differentially breaks through into awareness, generating a P3 component in the EEG. Meanwhile, if the suspected own-name was truly not familiar to the participant, it would not breakthrough into awareness and would be no different to the other names in the stream. The method has also been shown to be successful in detecting when participants perceived familiar (famous) faces among unfamiliar faces in RSVP streams (Alsufyani et al., 2019). A related method to the Fringe-P3 has also been used to detect concealed information of face, name, and word stimuli on the fringe of awareness based on involuntary eye movements (Rosenzweig and Bonneh, 2020). Perhaps most significantly, Bowman et al. (2014) also showed that the Fringe-P3 method was resistant to the countermeasure that has been argued to confound, or at the least, significantly complicate, ERP concealed knowledge tests (Rosenfeld, 2005; Rosenfeld et al., 2008). This countermeasure (which we called irrelevant as high salient) involves participants artificially elevating their response to the irrelevant, a control stimulus presented as frequently as the guilty knowledge (here, the participant's own name). This countermeasure is rendered ineffectual, because RSVP means participants are unable to (or at the least, significantly hampered in their efforts to) determine what the irrelevant is. As a result, they are unable to artificially elevate their response to it.

The Fringe-P3 method is a promising approach to deception detection. However, to assess the full generality of the method, it is important that a number of further empirical questions are answered. The two questions that we focus on here are as follows.

1) Can high detection rates be obtained with weaker name stimuli than "own-name"? In particular, Bowman et al. (2014, 2013) focused on "own-name" deception, i.e. participants attempted to conceal their real name and the Fringe-P3 system showed that they still exhibited a differential electrical brain response to it. Own-name is very special though, perhaps being the most overlearnt stimulus that can be presented to the brain. A classic demonstration of this is, of course, the Cocktail Party Effect (Cherry & Taylor, 1954). Others have also found faster and stronger responses to participants seeing their own-names than to other personally familiar or famous names (Mack & Rock, 1998; Yang et al., 2013). So, to counter the argument that a large electrical response was obtained in Bowman et al. (2014, 2013), specifically because own-name was the probe (guilty knowledge) stimulus, it is important to see whether a significant P3 is induced by name stimuli that are salient, but not as extremely so as own-name. In particular, in this respect, obtaining a significant group-level effect is not a massive challenge, but being able to demonstrate a robust effect at the individual level is not trivial.

2) Will the method work if we present multi-item information as the probe stimuli? In particular, Bowman et al. (2014, 2013) presented first names as the frames of RSVP streams. Thus, each frame was a single item and all frames were of the same kind: first names. However, some forms of guilty knowledge are intrinsically multi-item. For example, one might wish to demonstrate that a date (e.g. Eleventh May), a first-second name pair (e.g. Martin Jones) or an address part (e.g. Twenty Two High Street) is salient to a suspect. Therefore, it is important to evaluate whether the method will also work for such multi-item information as a step towards using it with real multi-item guilty knowledge.

This paper, then, provides evidence that both these issues can be addressed within the Fringe-P3 framework. Specifically, we consider these two issues as follows.

1) *Weaker Salience*: Famous names are a canonical example of stimuli that are salient to a broad population of individuals, but they would not be expected to be as salient as own-name. We assess whether famous names differentially break into awareness when presented in RSVP and whether we can detect these breakthrough events in Event Related Potentials (ERPs). In addition, we hypothesise that only items that are salient to a participant's perceptual system, such as the famous names, will be reportable in the recall and recognition tests at the end of the streams.

2) *Multi-item Stimuli*: Attentional blink experiments exhibit a phenomenon known as lag-1 sparing (Bowman & Wyble, 2007; Chun & Potter, 1995), whereby a second target is very accurately reported if it immediately follows a first target. Furthermore, this effect has been shown to generalise beyond two targets to sequences of three, four, or more consecutive targets (Olivers et al., 2007; Wyble et al., 2009), prompting an episodic theory of attention (Wyble et al., 2009). This suggests that multi-item stimuli might be effectively presented as a continuous sequence of RSVP frames. We investigate this hypothesis here by presenting first and second name pairs as consecutive frames, e.g. "Barack" followed immediately by "Obama".

This paper, then, demonstrates that the Fringe-P3 method is generalizable to weaker salience and to multi-item stimuli. We do this by demonstrating differential ERP responses to famous first and second name pairs, with such stimuli eliciting strong P3 components. Our first demonstration of this is at the group-level, but then we show that we can also detect this pattern on an individual basis. Such per-individual analysis is especially significant, since practical application of the method would be made on individuals, e.g. it is individuals that are found guilty, not populations.

2. Methods

2.1. Participants

Fifteen participants took part in the experiment, the same number as in our own-names experiment (Bowman et al., 2013). One participant was excluded due to a technical error with the recording system, leaving fourteen participants (aged 19-24, 6 male, 8 female) for analysis. All participants were students at the University of Kent, native English speakers, right-handed, free from neurological disorders, and had normal or corrected-to-normal vision. The study was advertised publicly, and each participant was paid 10 pounds (GBP) for participating. The Sciences Research Ethics Advisory Group at the University of Kent approved the study, and participants signed a consent form before participating in the experiment.

2.2. Stimulus Presentation

All stimuli were presented in RSVP using the Fringe-P3 method (Alsufyani et al., 2019; Bowman et al., 2013, 2014). RSVP streams were presented on a 20-inch LCD screen with a refresh rate of 60 Hz and a resolution of 1600x1200. The screen was placed at a distance of 60-80cm from the participant. Stimuli were presented using the Psychophysics Toolbox Version 3 and Matlab 2012. Stimuli were 16-point, light grey, monospaced, and sans-serif characters presented on a black background. The visual angle for each stimulus was 0.95° in height and 2.2° in width. The whole screen consisted of a rectangle of 28.52° by 37.56°. All items in the stream were presented at the same location on the screen. The Stimulus Onset Asynchrony (SOA) was 133ms, the same as Bowman et al. (2013) and Alsufyani et al. (2019).

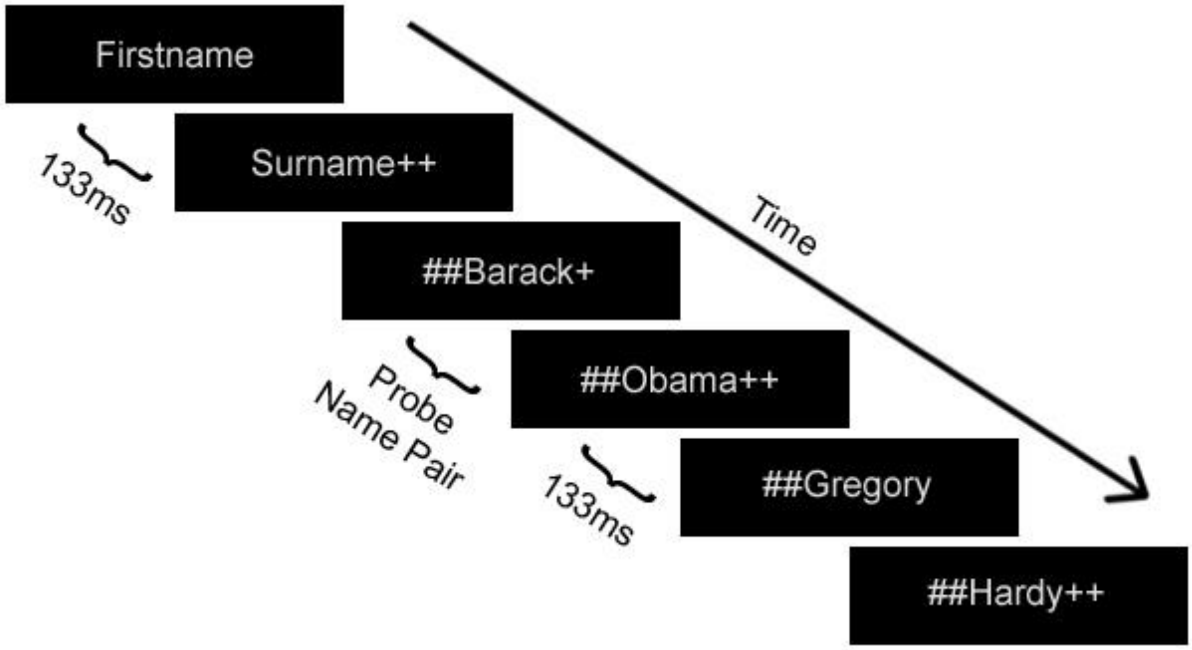

Figure 1. Example Section of an RSVP Stream of Name Stimuli.

Each RSVP stream consisted of 15 pairs of first and second names, one of which was the critical pair (target, probe, or irrelevant) while the rest were distractors. Streams had an SOA of 133ms. The critical name pair in this example is the probe name pair Barack Obama.

2.3. Stimuli and Experimental Design

All stimuli were names presented as pairs of first and second names. The names in the pair were presented in sequence, first name followed by second name, similar to Proverbio et al. (2009). An example of a section from an RSVP stream of name stimuli is shown in figure 1. Each RSVP stream consisted of 15 pairs of first and second names, one of which was a critical pair. There were three categories of critical pair stimuli (probe, irrelevant, and target) with 5 name pairs in each category. The probe stimuli were names of famous people from different fields such as politics, entertainment, and sport: Justin Bieber, Leonardo DiCaprio, Barack Obama, David Beckham, and William Shakespeare. Due to a programming error, three participants (2, 3, and 4) were presented with Winston Churchill instead of

William Shakespeare, and one participant (4) was also presented with Britney Spears instead of David Beckham. We do not believe that this error affected participants' results and further discussion of this can be found in Appendix A. The irrelevants were random names that were unknown to the participants: Belia Labbe, Audrey Slater, Annie Rand, Blyth Tomayo, and Kylie Carr. The target was a single pair of names that participants were instructed to respond to. At the start of the experiment, participants were presented with a list of five name pairs and asked to select their target name pair, similar to Bowman et al. (2014). Three participants (1, 5, and 6) chose Katherine Stevenson, two participants (9 and 12) chose Christopher Steffen, one participant (11) chose Waylon Travis, and the remaining eight chose April Harrington. No participant chose Waleed Finch. The target stimulus was included just to keep participants attention on the stream and plays no role in the analyses we perform.

The experiment was divided into 5 blocks, each comprised of 15 probe trials, 15 irrelevant trials, and 15 target trials (45 trials total per block). Each block used one probe name pair and one irrelevant name pair, which were presented 15 times each. The same target was used for every block and presented 15 times in each block. In total there were 75 probe trials, 75 irrelevant trials, and 75 target trials (225 trials overall) in the whole experiment. The order in which the trials appeared within each block was randomised.

Each trial consisted of one critical name pair (probe, irrelevant, or target), and 14 random name pairs used as distractors. The position of the critical name pair within the stream was selected pseudo-randomly, so that it had an equal probability of appearing in the 5th position through to the 10th position. These positions were chosen to prevent any beginning or end of stream effects overlapping with the ERP response to the critical item. Each trial began with a fixation stimulus (XXXXXXXXX) presented for 800ms to position the participant's focus on the stimulus presentation area. At the end of each trial, either "--------- " or "=========" was presented (randomly selected) for 133ms. Participants were then asked to report what the last item was, in order to maintain participant's attention for the whole length of the stream.

Distractors were selected randomly from a database containing 10,000 pairs of common first and second names (50% male and 50% female). This database was populated using a tool that generates random first and last names, based on the US Census database (http://random-name-generator.info/). The irrelevant and target names were also chosen using this tool. All names were presented with a maximum length of 11 characters, and the first letter capitalised. To ensure an equal length of all presented names, shorter ones were padded using a randomising algorithm, with '#' or '+' characters randomly blocked on either side of the word (e.g. '###Alice+++', '###Gregory+', 'Danielle+++'). This fits with the presentation of name stimuli in Bowman et al's Fringe-P3 papers (2014, 2013). Distractors for each stream were chosen from the database pseudo-randomly. In order to avoid repetition, names could not contain two or more letters in the same position as their immediate predecessor in the stream. In addition, names that shared three or more letters, in the same position as one of the critical items, were not presented as distractors.

2.4. Tasks

Before the experiment, participants selected a target name to look for in the streams. At the end of each stream they were asked what the last item in the stream was, as described above, followed by "Did you see the target name?" Participants responded using a numeric keypad with their right hand, pressing '1' for Yes and '2' for No. The target task was included to ensure participants paid attention to the streams. At the end of each block, there were also recall and recognition tests, which are explained in the next section.

Participants were instructed to keep their eyes fixated on the centre of the screen, and to avoid eye and body movements during each trial. Participants were not told in advance that there would be famous names within the streams.

Before the EEG recording, there was a training session of 20 practice trials followed by a recall and a recognition test. The aim of this practice session was only to familiarize participants with the presentation of the stimuli and to make sure that they could identify the target. These practice trials only contained distractors and the target and did not contain any of the probe names from the main experiment or the names of any other famous people.

2.5. Recall and Recognition Tests

Both tests were conducted to explore participants' memory and were performed at the end of each block. For the recall test, participants were asked on screen "What did you see?" and were instructed to use the keyboard to write any names that they saw during that block. There was no limit to the number of names that they could enter, and they were encouraged to recall as many names as possible.

For the recognition test, participants were presented with five categories of name pairs (first name – second name) and asked to give a confidence rating of how often each one of them appeared. The ratings were on a scale of 1 to 5, where 1 means 'Not appeared at all' and 5 means 'Appeared very often'. The five categories of name pairs that we presented were: target, probe, irrelevant, unpresented famous name (the name of a famous person that was not presented in streams during the experiment), and a distractor. The distractor names were chosen randomly from our distractors database of 10,000 name pairs and had the same very small probability (0.14 %) of being presented in a trial as any of the other possible distractor names in the database. All participants were asked about the same distractor names in the recognition test. There were five name pairs in each category – one for each block – excluding the target, of which there was only one for the whole experiment.

Recall is the more demanding test since memory retrieval is not cued with an item to search for. The recall task must be performed before the recognition task, otherwise it would not be free recall as the recognition test would reveal names to recall.

2.6. Data Acquisition

We recorded EEG data using a BioSemi ActiveTwo system (BioSemi, Amsterdam, The Netherlands). The data were digitized at 2048Hz during acquisition and were filtered with a low-pass of 100Hz at the time of recording. Impedances were kept below 10 kOhms. Electroencephalographic (EEG) data were recorded at the Fz, Cz, Pz, P3, P4, Oz, A1, and A2 electrodes based on the standard 10-20 system. During recording, data were referenced to a ground formed from a common mode sense (CMS) active electrode and driven right leg (DRL) passive electrode. The electrooculograms (EOG) generated from blinks and eye movements were recorded from the participant’s left and right eyes, using two bipolar HEOG and VEOG electrodes.

2.7. Analysis Procedure

The electroencephalographic data were analysed using EEGLAB version 13.6.5b and Matlab 2016a. The data were re-referenced to the average of the combined mastoids (electrodes A1 and A2) and were resampled to 512Hz. The data were then filtered with a low-pass of

45Hz and high-pass of 0.5 Hz. In order to remove the Steady State Visually Evoked Potentials (SSVEP) oscillation set-up by the stream presentation, notch filters were applied between 7 and 9Hz. The data were then epoched into segments using a -100ms to 1200ms stimulus-locked window. All epochs were time-locked to the onset of the first name of a critical pair. ERPs for the target, probe, and irrelevant were generated by averaging all trials in each condition (separately).

Eye blinks were detected and rejected for activity below -100 µV or above +100 µV in the EOG channels. For the scalp channels, trials containing electrical activity below -50 µV or above +50 µV were rejected. Baseline correction was then performed from -100 to 0ms. The number of trials remaining after artefact rejection, per condition, ranged between 62 and 75 for probe (*M* = 71.79, *SD* = 3.60) and irrelevant (*M* = 71.57, *SD* = 4.22). The maximum possible trials were 75 for each condition. No further participants were excluded from the study following artefact rejection.

2.8. ERP Analysis Procedure

The ERP data were analysed using the Aggregated Grand Average of Trials (AGAT) method (Brooks et al., 2017, Bowman et al., 2020). Data from the Pz electrode were analysed, as the P3 is typically maximal from that electrode (Comerchero & Polich, 1999; Polich & Kok, 1995). The probe and irrelevant ERP and behavioural data were the focus of all analyses. The target data was not analysed, since the target was included simply as a task to keep participants' attention on the streams and would provide no interesting data in regard to the detection of famous names.

The AGAT uses an orthogonal contrast between probe and irrelevant ERPs by choosing a window of interest based on the average of all probe and irrelevant trials in one aggregated

ERP. By using the aggregated ERP, neither probe nor irrelevant trials are given biased treatment, thus preventing inflation of the false positive (type I error) rate that could happen if the window of interest was chosen based purely on the probe ERP, see Friston et al. (2006), Brooks et al. (2017), and Bowman et al. (2020) for further justification of this approach.

The AGAT also uses mean amplitude measurements, which have been shown to be more robust against high frequency noise (Luck, 2005) and have been used in previous studies to measure the P3 for familiar and unfamiliar names (Pickering & Schweinberger, 2003; Schweinberger et al., 2002; Tacikowski et al., 2011).

2.8.1. Group Level

The aggregated ERP was created by averaging all probe and irrelevant trials into one ERP. An algorithm then searched along this ERP between 300ms and 1000ms for the 100ms window with the highest mean amplitude (window of interest). These timings and window sizes were chosen based on previous studies using the Fringe-P3 method (Alsufyani et al., 2019; Bowman et al., 2013, 2014). This window of interest was then applied separately to the probe and irrelevant ERPs and the mean amplitudes within that time frame calculated for both. The irrelevant mean was then subtracted from the probe mean to give the true observed difference to be used in further analysis.

For the group level analysis, the window of interest and mean amplitudes were found for each participant and these means then underwent a two-tailed paired t-test.

2.8.2. Individual Level

The individual level analysis uses the same method as the group level up to and including the calculation of the true observed value, but then performs a Monte Carlo permutation test instead of a t-test. Under the null hypothesis, the irrelevant and probe trials are samples from the same distribution – the null distribution – and thus would be exchangeable during permutation. Permutation tests have been shown to be the preferred null hypothesis test for ERP data, since bootstrapping has been shown to be biased with peak measures while permutation is not (Zoumpoulaki et al., 2015).

After finding the true observed value for the participant, the individual probe and irrelevant trials were permuted to create surrogate probe and irrelevant ERPs. The window of interest found earlier was then applied to these surrogate ERPs and their mean amplitudes calculated and subtracted to find the surrogate mean difference. This was repeated 10,000 times until there were 10,000 surrogate differences. The p-value was then calculated as the proportion of surrogate differences that were larger than the true observed difference.

3. Results

3.1. Group Level Analyses

Table 1 shows the number of times participants answered "yes" when asked at the end of the trial if they saw the target in the target, probe, and irrelevant conditions. This shows that participants correctly said yes after most target trials and, with the exception of participant 2, (who may have misunderstood the instructions), rarely said yes (7 or less times) after probe or irrelevant trials.

Table 1. The number of times participants answered "yes" when asked if they saw the target in each condition

Participant	*Target*	*Probe*	*Irrelevant*
1	62	1	2
2	69	30	19
3	52	2	3
4	70	6	6
5	65	4	5
6	56	7	5
7	73	1	1
8	60	2	3
9	48	6	1
10	74	1	4
11	48	7	2
12	65	5	2
13	66	5	4
14	64	7	3
Mean	62.286	6	4.286
Median	64.5	5	3
Std Dev	8.534	7.296	4.497

This table shows the number of times participants said “yes” after target trials (hits), and probe and irrelevant trials (false alarms). 10/14 participants scored at least 60 hits. The maximum possible score for each condition was 75.

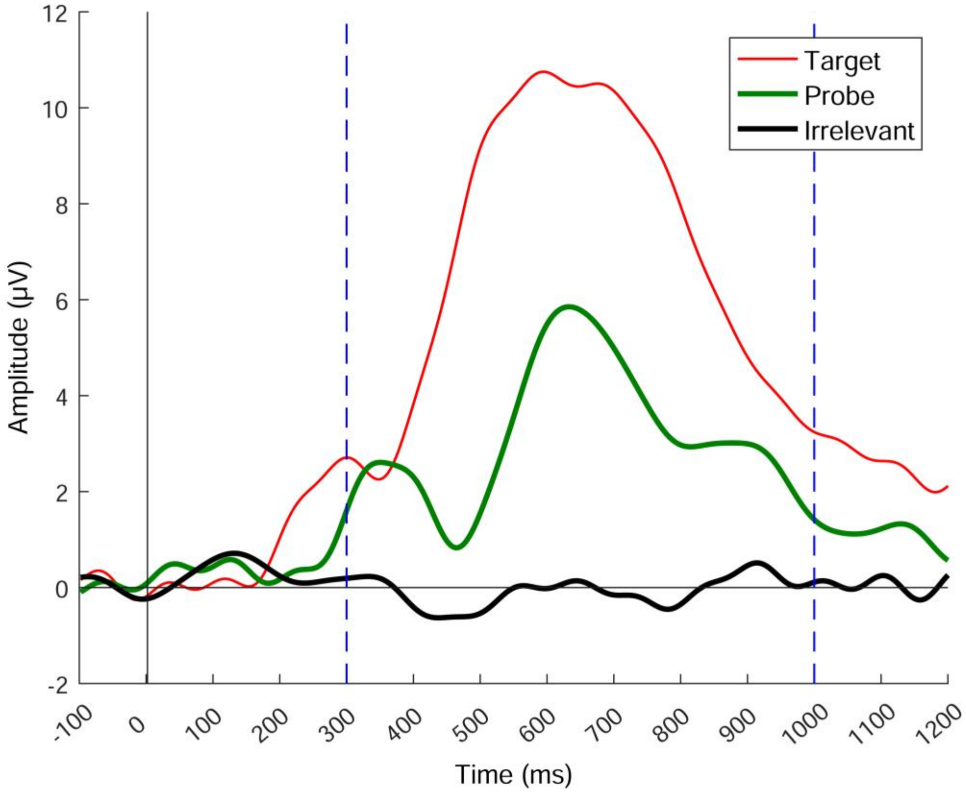

Figure 1. Grand Average for All Participants’ Probe and Irrelevant from the Pz Electrode.

All ERPs were time locked to the onset of the first name in a pair. The blue dashed line marks the start and end of the AGAT search window. Extra smoothing with a low pass filter of 10 Hz was used in this figure, purely for presentational purposes, and was not applied to the time-series analysed. Both the target and probe elicited a large P3, but not the irrelevant. As the target is task-relevant, it was expected to elicit the largest P3. However, the comparison we are interested in is between probe and irrelevant, with no involvement of the target.

It can be observed in figure 2, that the target and probe elicited large P3s at Pz within the 300ms to 1000ms time window, but the irrelevant did not. The mean amplitude values (see

Table 2) found in the AGAT analysis were higher for the probe compared to the irrelevant. A bar graph of the group mean amplitudes for each condition can be found in figure 3. A paired t-test was conducted on the mean amplitude values for probe and irrelevant for all participants and found a highly significant difference between the probe (*M* = 6.077, *SD* = 1.816, *Mdn* = 6.497) and the irrelevant (*M* = 0.372, *SD* = 1.321, *Mdn* = 0.114), $t(13) = 9.668$, $p < 0.0001$, and a very large effect size (Cohen's $d = 3.593$). The ERPs and investigation of a P3a pattern at Fz and Cz can be found in appendix B.

Table 2. Individual Participants' Mean Amplitudes and P-Values from the Pz Electrode.

Participant	*Probe*	*Irrelevant*	*P-Value*
1	7.58	1.23	0.0003
2	4.64	0.27	0.0379
3	7.63	-1.44	0.0002
4	5.86	2.56	0.0734
5	6.80	-1.75	<0.0001
6	6.19	-0.32	<0.0001
7	8.18	0.64	<0.0001
8	3.01	1.26	0.1889
9	2.61	-0.72	0.025
10	8.08	2.27	0.0109
11	4.52	-0.32	0.0124
12	6.89	-0.28	<0.0001
13	7.58	-0.04	0.0019
14	5.51	1.85	0.0195
Mean	6.077	0.372	0.026
Median	6.497	0.114	0.006
Std Dev	1.816	1.321	

This table presents the mean and median amplitude values from the AGAT analysis for probe and irrelevant as well as the p-values from the individual participants level analysis. The mean and median amplitude values are larger for the probe than the irrelevant for all participants. 12/14 participants have p-values below 0.05. 6/12 participants have p-values below 0.001. The smallest p-value that can be obtained from the 10,000 permutations is 0.0001.

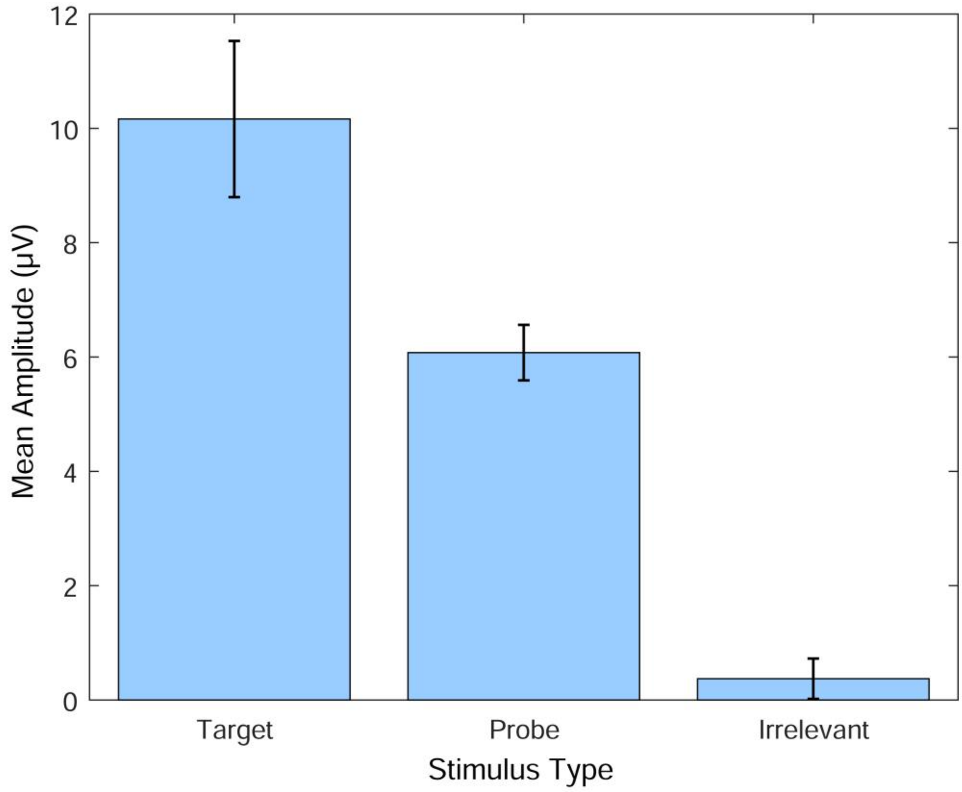

Figure 3. Group Mean Amplitudes of Target, Probe, and Irrelevant

This bar graph shows the grand average of the mean amplitudes found within each participant's window of interest for each condition. The same window of interest used for each participant's probe and irrelevant ERPs in the AGAT analysis was used for their target ERP to generate this graph. There was a highly significant difference between probe and irrelevant at the group level, $t\ (13) = 9.668$, $p < 0.0001$, and a very large effect size, Cohen's $d = 3.593$. The target was not statistically analysed.

3.2. Individual Level Analyses

The strong group effect was supported by the individual level analyses. 86% of all participants (12/14) showed a significant difference between the probe and irrelevant mean amplitudes at an alpha level of 0.05. Six participants (43%) have p-values below 0.001 and six obtained p-values between 0.001 and 0.05. Individual p-values are presented in Table 2.

The mean p-value was 0.026 and the median was 0.006. The median is the most suitable measure of central tendency here, as the distribution of p-values is skewed.

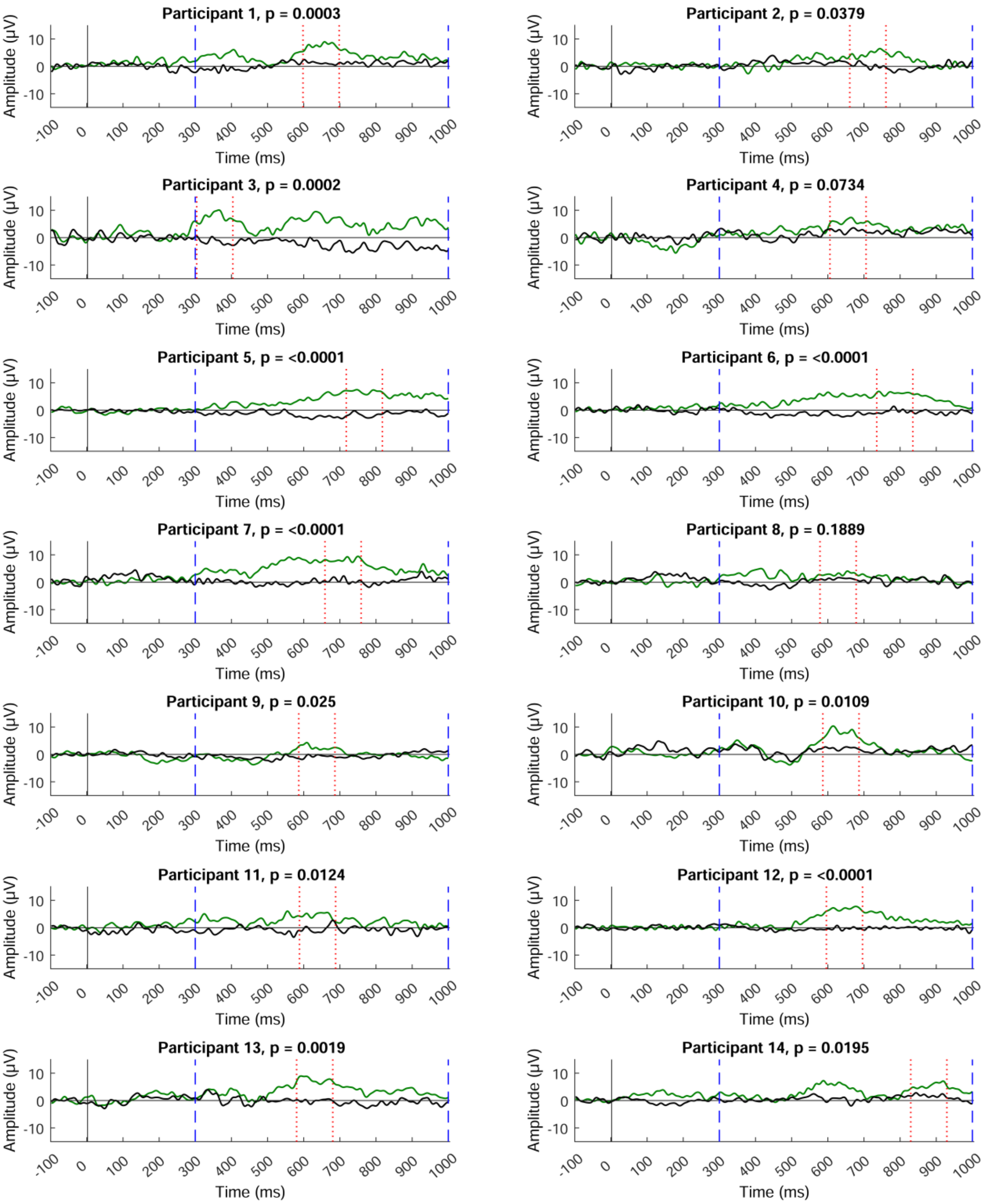

Figure 4. Individual Participants' ERPs from the Pz Electrode.

Each ERP is labelled with the corresponding participant number and p-value. The green line represents the probe ERP and the black line represents the irrelevant ERP. The dashed vertical blue lines mark the start and end of the AGAT search window. The dotted vertical red lines represent the start and end of the window of interest with the highest mean amplitude used in the analysis for that participant.

Figure 4 shows individual participants' probe and irrelevant ERPs. As can be seen, most of the participants showed clear positive peaks for probes within the expected 300ms to 1000ms window. Only two participants (4 and 8) did not have significant differences between their probe and irrelevant mean amplitudes in the analysis. Participant 4 showed a positivity around 650ms for the probe, and the AGAT correctly selected this as the window of interest around this positivity, but there was not enough difference between the probe and irrelevant mean amplitudes to outweigh the (background) variability in the data and reach significance (p = 0.073). This is most likely due to the high noise level in the data causing the irrelevant to be positive in that window.

Participant 8 showed a positivity for the probe peaking at around 400ms, but the irrelevant was negative at this point, causing that region to have a smaller amplitude in the aggregated ERP and not be selected in the AGAT analysis. The algorithm instead picked a later region with a smaller probe positivity, resulting in a smaller difference between the probe and irrelevant and a high p-value (p = 0.189). The irrelevant being negative during the earlier probe positivity is likely to be due to noise in the data. In contrast, participant 12, one of the four participants with the strongest significant differences, showed a high positive peak which resulted in a high probe – irrelevant mean amplitude difference and a highly significant p-value (p < 0.0001). Null hypothesis distributions for participants 8 and 12 are presented in figure 5. The null hypothesis distribution for participant 8 shows that the true observed value (probe – irrelevant mean amplitude) could not reject the null hypothesis, which led to a high p-value; $p = 0.189$. On the other hand, the true observed value for participant 12 falls well outside the null hypothesis distribution, resulting in a highly significant p-value ($p < 0.0001$).

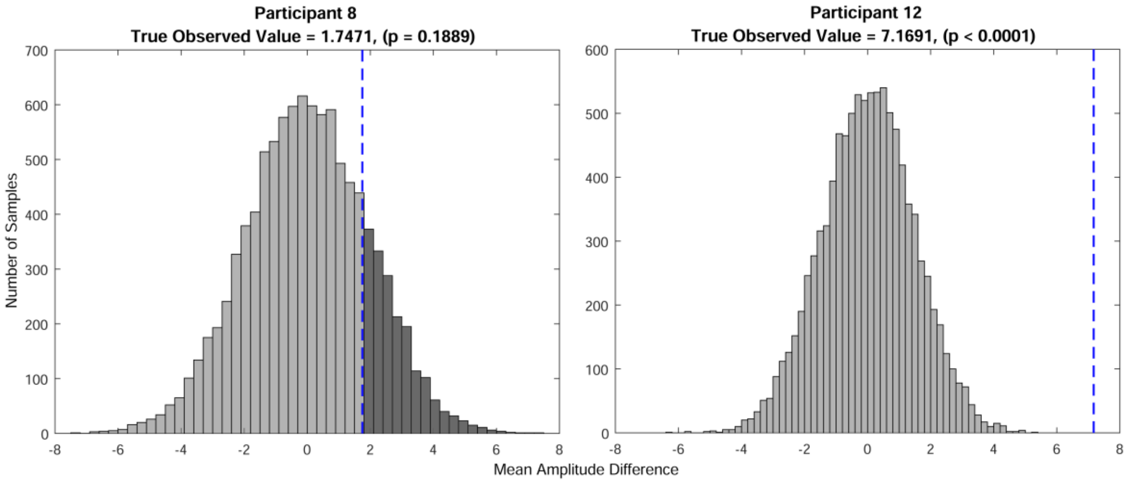

Figure 5. Null Hypothesis Distributions for Participants 8 and 12.

This figure shows the probe - irrelevant mean amplitude differences from the permutation analyses for participants 8 and 12. The dashed vertical lines represent the true observed value.

3.3. Recall and Recognition Tests

Individual participants' recall scores are presented in Table 3 and a bar graph of the mean recall scores is in figure 6. There were five probes and five irrelevants used in the experiment, so total recall scores are out of five. The results showed that all fourteen participants were able to recall two or more of the probes. Thirteen of the participants (93%), recalled at least three probes. In contrast, only two participants were able to report any of the irrelevants, both of whom recalled only a single irrelevant. A paired samples t-test found a significant difference between probe recall scores (*M* = 4.214, *SD* = 0.975, *Mdn*

= 4.5) and irrelevant recall scores (*M* = 0.143, *SD* = 0.363, *Mdn* = 0), $t\ (13) = 13.350$, $p < 0.0001$, and a very large effect size (Cohen's $d = 5.534$). The recall measure was found not to be normally distributed, so a Wilcoxon's signed-rank test was performed in addition and also found a highly significant difference between probe and irrelevant recall scores, $Z = -3.347$, $p < 0.001$ (approx.).

In the recognition test, participants were instructed to give a confidence rating on a scale of 1 to 5 of how often a particular name pair was presented in that block. Five categories of name pairs were presented: target, probe, irrelevant, un-presented famous, and a distractor. The means of the ratings across the five blocks were used as the participants' final confidence ratings and are presented in Table 4. Figure 7 presents a bar graph of the mean final confidence ratings. The main comparisons were probe (*M* = 4, *SD* = 0.532, *Mdn* = 4) against un-presented famous name (*M* = 1.286, *SD* = 0.321, *Mdn* = 1.2), and irrelevant (*M* = 1.586, *SD* = 0.454, *Mdn* = 1.5) against the distractor name (*M* = 1.50, *SD* = 0.280, *Mdn* = 1.5). Paired t-tests found a highly significant difference between the probe and un-presented famous name confidence ratings, $t\ (13) = 16.480$, $p < 0.0001$, and a very large effect size (Cohen's $d = 6.179$). In contrast, there was no significant difference between the irrelevant and the distractor, $t\ (13) = 0.763$, $p = 0.459$, and a small effect size (Cohen's $d = 0.227$), suggesting that participants were able to reliably recognise the probe names but not the irrelevant names, despite the irrelevants being repeated just as often.

Table 3. Individual Participants' Recall Test Scores for Probe and Irrelevant.

Participant	*Probe*	*Irrelevant*
1	2	0
2	5	0
3	5	0
4	4	0
5	4	1
6	5	0
7	5	0
8	4	0
9	3	1
10	3	0
11	5	0
12	5	0
13	4	0
14	5	0
Mean	4.214	0.143
Median	4.5	0
Std Dev	0.975	0.363

This table presents the total number of probe and irrelevant names recalled by each participant during the recall tests at the end of each block. The highest possible score is 5. All participants recalled at least two probes and 13/14 recalled at least three probes. Only two participants recalled a single irrelevant.

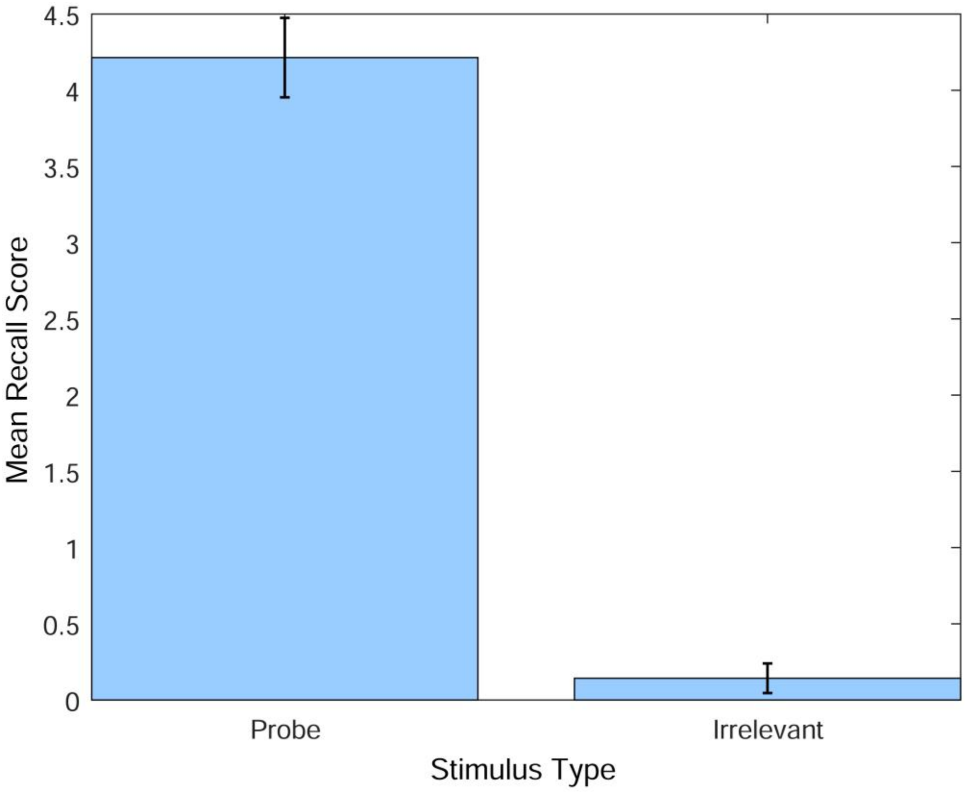

Figure 6. Mean Recall Scores.

This bar graph shows the mean recall scores for probe and irrelevant. The highest score possible was 5. The mean probe score was significantly higher than the irrelevant score.

Table 4. Individual Participants' Confidence Ratings from the Recognition Tests.

Participant	*Probe*	*Irrelevant*	*Unpresented Famous*	*Distractor*
1	3.8	1	1	1
2	3.8	2.4	2	1.8
3	4	1.6	1	1.6
4	4.4	1.8	1.4	1.2
5	4.2	2	1.6	1.6
6	4	1.2	1	1.8
7	4.8	1.4	1.4	1.4
8	3.4	1.2	1.2	1.4
9	3.4	2.2	1.8	1.8
10	4	2	1.2	1.8
11	4	1.4	1	1.2
12	5	1	1.2	1.4
13	4.2	1.8	1.2	1.2
14	3	1.2	1	1.8
Mean	4	1.586	1.286	1.5
Median	4	1.5	1.2	1.5
Std Dev	0.532	0.454	0.321	0.280

This table shows the final confidence ratings in each category for all participants. These ratings are the means of the ratings given at the end of the five blocks, where 1 meant the participant thought the name pair did not appear at all, and 5 meant the name pair appeared very often.

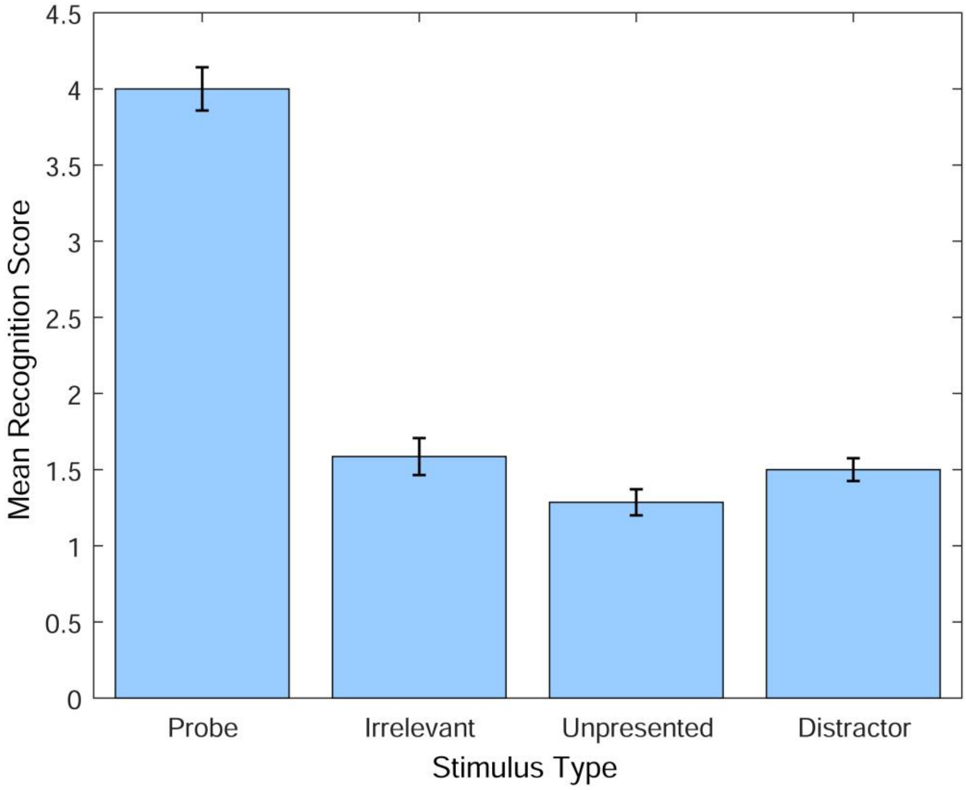

Figure 7. Mean Confidence Ratings from the Recognition Tests.

This bar graph shows the means of the final confidence ratings from the recognition tests. The highest possible score was 5. The mean probe rating was significantly higher than the irrelevant, unpresented, and distractor mean ratings.

4. Discussion

Bowman et al. (2013, 2014) demonstrated that presenting stimuli on the fringe of awareness using RSVP provides an effective method to detect identity deception. They showed that salient stimuli (i.e. a participant's own first name) in RSVP broke through into awareness, resulting in the generation of a P3 component, which could reliably be used to differentiate between those who were attempting to conceal their true identity versus those who were not. Their Fringe-P3 studies (Bowman et al., 2013, 2014) demonstrated a very high hit rate (correctly detecting deception) and a low false alarm rate (incorrectly detecting deception when there was none) using P3 components. However, due to the high classification accuracy that they obtained, a question might be raised as to whether the use of own-name stimuli is the main reason for obtaining such high accuracy. Therefore, the first aim of the present study was to investigate the effectiveness of the Fringe-P3 approach with name stimuli that are still salient, but less so than own-name. Thus, we explored whether famous names presented on the fringe of awareness, differentially breakthrough into consciousness, and whether we can detect these breakthrough events with ERPs. The strong P3 and significant results at the group level and for 86% of participants at the individual level demonstrates that the Fringe-P3 method can successfully detect when weaker (famous) name stimuli breakthrough into awareness. This finding is in line with other studies that have used famous names as stimuli without RSVP and found that famous names elicited a P3 (Pickering & Schweinberger, 2003; Proverbio et al., 2009; Schweinberger et al., 2002; Stenberg et al., 2009; Tacikowski et al., 2011; Tacikowski & Nowicka, 2010). The additional recall and recognition tests also showed that participants were able to reliably recall and recognise the probe famous names but not the irrelevant or distractor names, as expected, thus providing further evidence that only the salient names stood out from the RSVP streams and broke through into awareness.

The second aim of this experiment was to see if multi-part stimuli could be successfully used with the Fringe-P3 method. Specifically, we presented the first name followed by the second name in adjacent frames in the RSVP stream. This generated a robust P3 pattern for famous name pairs and significant p-values, showing that the Fringe-P3 method can indeed be used effectively with multi-item stimuli. This finding is consistent with theories of the attentional blink (Bowman & Wyble, 2007; Hommel & Akyürek, 2005; Wyble et al., 2009), which emphasize that two consecutive targets (lag-1) are processed together (Simione et al., 2017; Wyble et al., 2009), with high accuracy for both, and particularly the second. The spreading the sparing phenomenon (Olivers et al., 2007) and theories of episodic attentional processes (Wyble et al., 2009) suggest that even longer sequences of consecutive salient items will not generate an attentional blink and will thus also be well perceived in RSVP, opening up the possibility that longer series of identifying information, such as addresses, might be included in the Fringe-P3 method. It is worth noting that there is also the possibility that the first and second name pairs could have been perceived as a single integrated percept in the fashion introduced in Hommel & Akyürek (2005) and Akyürek et al (2012). This possibility would not, however, impact the inferences we are making in this paper, which simply concern the effectiveness of the Fringe-P3 method to detect P3s for multi-item stimuli.

This experiment and the one reported in Alsufyani et al. (2019) have also enabled us to assess the effectiveness of the Fringe-P3 method in two further respects. Firstly, the use of fame-related stimuli (names here and faces in Alsufyani et al. (2019)) has enabled us to explore the effectiveness of the method to a purer form of intrinsic salience. That is, in this experiment, there was no explicit task associated with famous names, and participants were not informed of their presence in advance (as was the case for famous faces in Alsufyani et al. (2019)). This enabled us to see whether, even when there is no explicit task associated with them, the brain still selects salient stimuli. Secondly, the previous own-name Fringe-P3 demonstrations (Bowman et al., 2013, 2014) have effectively been detection tasks for the participant's perceptual system (i.e. it was searching for an item – their own name). In contrast, once participants have determined that fame-related stimuli are being presented, which they will most likely do at some point in our two fame experiments, they may start (at least implicitly) "looking" for items in this category. Consequently, this effectively becomes an implicit identification task; one in which participants' brains are seeking to identify the fame-related stimuli presented. Thus, the findings in this paper provide evidence that the Fringe-P3 method is also effective at determining stimuli that are, firstly, intrinsically salient (but not initially associated with an explicit task) and, secondly, being searched for in an effectively implicit identification task.

With respect to the P3 shape, the P3 elicited by the famous names (from around 400ms to 1000ms) is much broader in time than the corresponding P3 elicited by own-names in previous Fringe-P3 experiments; 300-600ms in Bowman et al. (2013) and 300-780ms in Bowman et al. (2014). This is likely to be due to the famous names being presented in two frames (first followed by second name). Thus, in a sense, there were two salient names presented in immediate succession. The grand average ERPs in figure 2 show that the P3 probe does not appear to consist of two individual P3s for first names and second names. This suggests that the first and second name proceeded into working memory together, resulting in a longer single P3. This finding is in line with Craston et al.'s (2009) attentional blink investigations, which reported a single combined long P3 pattern for lag-1 where T2 immediately follows T1; see also Jones et al. (2020) and Pincham et al. (2016) for further discussion of extended P3s.

In addition, this longer P3 component justifies the choice of a mean amplitude measure for analysing the P3 in the present study, instead of the Peak-to-Peak, which was used in previous Fringe-P3 studies (Bowman et al., 2013, 2014). The Peak-to-Peak method is based on finding the difference between the maximum positivity and (the following) negativity within a specific ERP time window. Therefore, the use of the Peak-to-Peak by Bowman et al. (2013, 2014) seems reasonable as their P3 component indeed consisted of a positive peak followed by a negative deflection, within a time window from 300ms to 1000ms, with respect to the stimulus onset. As such a negative deflection was absent from the ERPs in the present study, the Peak-to-Peak measure would be inappropriate for detecting the P3 here. A study by Bergström, Anderson, Buda, Simons, & Richardson-Klavehn (2013) has shown that Peak-to-Peak performed worse than a mean amplitude measure with ERPs that did not contain clear post-P3 negative peaks. Moreover, the use of the mean amplitude measure in this study is consistent with Alsufyani et al. (2019) who used similar measures for analysing the Fringe-P3 method with famous faces, and with Craston et al. (2009) who used the same measure for analysing a joint P3 component of T1 and T2 at lag-1.

Although this experiment is not a true concealed information paradigm, it is a step towards testing real guilty knowledge, therefore, one potential criticism of it is the lack of an innocents group to ensure there is no inflation of the false positive. In the current experiment there is always a salient probe for each participant, but in an innocents experiment there would not – i.e. what the person administering the test refers to as the probe would actually be an irrelevant to the innocent participant since it would not be salient. The current experiment and previous Fringe-P3 work shows that participants do not generate P3s for irrelevants. Therefore, this suggests that there should not be an inflation of false positives, i.e. innocents mistakenly adjudged to be guilty, in a concealed information version of the paradigm presented here.

Additionally, we ran an innocents test in the 2014 paper and found no P3s or inflation of the false-positive rate for the irrelevants/innocent probes. This is consistent with our other experiments on repeating non-salient stimuli across RSVP streams. These found that the probability of a participant “seeing” a repeating non-salient item does not increase with repeated presentations (Avilés et al., 2020) and found no evidence that participants can differentially “see” non-salient repeating items even when their task is to “look for them” (Bowman et al., 2014). Therefore, it seems unlikely that innocent participants would see a repeating non-salient item and result in false detection when they are not even looking for it.

In summary, the findings reported here provide evidence that the Fringe-P3 approach is successful in detecting when participants perceived name stimuli that are familiar but have weaker salience than own-name stimuli, and that the method is effective with multi-item stimuli. Significant differences were found between famous name (probe) and unfamiliar name (irrelevant) ERPs at both the group and individual participant levels. The use of famous names in this study has shown that broad familiarity is sufficient to mark a stimulus out as salient in RSVP subliminal search, resulting in the generation of a P3 component. When placed within the context of Bowman et al. (2013, 2014) and Alsufyani et al. (2019), the results here further demonstrate the potential of the Fringe-P3 method in forensic applications that require the use of different types of stimuli beyond own first name stimuli, including multi-item stimuli, such as full names, dates, and street addresses. Within forensic settings, the findings in this paper are particularly relevant to a set of identity related applications. Most significantly, the results, especially the accuracy in per-individual analyses, suggest that the Fringe-P3 method could be used to detect whether a suspect has knowledge of a particular person. In this case, if the suspect was familiar with a particular person (guilty), their name would be salient and would breakthrough into awareness and generate a P3 component. If they were truly not familiar with the person (innocent), their name would not be salient, would remain subliminal, and would not generate a P3 or have a significant difference to the irrelevant. Demonstrating familiarity in this way could be used to identify criminals on the basis of who they know, such as victims or accomplices. The method could also potentially be applied to line-ups, especially since we have demonstrated the effectiveness of the method with faces, although currently, only for famous faces. This is a line of research we are pursuing.

Another aspect of future work will be to ensure that our countering countermeasures experiments, which have been performed with own-name probes, carry over to familiar names (with famous names being a particular example). We believe that our key demonstration that irrelevants cannot be searched for in RSVP to artificially make them salient and confound deception detection (Bowman et al., 2014) should still obtain. This is because the inability to search for repeating items seems to be inherent to rapid serial visual presentation in general, not the particular RSVP stimuli used (Avilés et al., 2020).

References

Acqualagna, L., & Blankertz, B. (2013). Gaze-independent BCI-spelling using rapid serial visual presentation (RSVP). *Clinical Neurophysiology*, *124*(5), 901–908. https://doi.org/10.1016/j.clinph.2012.12.050

Akyürek, E. G., Eshuis, S. A. H., Nieuwenstein, M. R., Saija, J. D., Başkent, D., & Hommel, B. (2012). Temporal target integration underlies performance at Lag 1 in the attentional blink. Journal of Experimental Psychology: Human Perception and Performance, 38(6), 1448–1464. https://doi.org/10.1037/a0027610

Alsufyani, A., Hajilou, O., Zoumpoulaki, A., Filetti, M., Alsufyani, H., Solomon, C. J., Gibson, S. J., Alroobaea, R., & Bowman, H. (2019). Breakthrough percepts of famous faces. *Psychophysiology*, *56*(1), e13279. https://doi.org/10.1111/psyp.13279

Avilés, A., Bowman, H., & Wyble, B. (2020). On the limits of evidence accumulation of the preconscious percept. *Cognition, 195*, 104080. https://doi.org/10.1016/j.cognition.2019.104080

Bergström, Z. M., Anderson, M. C., Buda, M., Simons, J. S., & Richardson-Klavehn, A. (2013). Intentional retrieval suppression can conceal guilty knowledge in ERP memory detection tests. *Biological Psychology*, *94*(1), 1–11. https://doi.org/10.1016/j.biopsycho.2013.04.012

Bowman, H., & Avilés, A. (n.d.). *Fragile Memories for Fleeting Percepts*. Manuscript in Preparation.

Bowman, H., Brooks, J. L., Hajilou, O., Zoumpoulaki, A., & Litvak, V. (2020). *Breaking the Circularity in Circular Analyses: Simulations and Formal Treatment of the Flattened Average Approach*. PLoS Computational Biology.

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. https://doi.org/10.1371/journal.pone.0090595

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*, *8*(1), e54258. https://doi.org/10.1371/journal.pone.0054258

Bowman, H., & Wyble, B. (2007). The simultaneous type, serial token model of temporal attention and working memory. *Psychological Review*, *114*(1), 38–70. https://doi.org/10.1037/0033-295X.114.1.38

Broadbent, D., & Broadbent, M. H. (1986). Encoding speed of visual features and the occurrence of illusory conjunctions. *Perception*, *15*(5), 515–524. https://doi.org/10.1068/p150515

Brooks, J. L., Zoumpoulaki, A., & Bowman, H. (2017). Data-driven region-of-interest selection without inflating Type I error rate. *Psychophysiology*, *54*(1), 100–113. https://doi.org/10.1111/psyp.12682

Chennu, S., Alsufyani, A., Filetti, M., Owen, A. M., & Bowman, H. (2013). The cost of space independence in P300-BCI spellers. *Journal of NeuroEngineering and Rehabilitation*, *10*(1), 82. https://doi.org/10.1186/1743-0003-10-82

Cherry, E. C., & Taylor, W. K. (1954). Some Further Experiments upon the Recognition of Speech, with One and with Two Ears. In *Journal of the Acoustical Society of America* (Vol. 26, Issue 4, pp. 554–559). https://doi.org/10.1121/1.1907373

Chun, M. M., & Potter, M. C. (1995). A Two-Stage Model for Multiple Target Detection in Rapid Serial Visual Presentation. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(1), 109–127. https://doi.org/10.1037/0096-1523.21.1.109

Comerchero, M. D., & Polich, J. (1999). P3a and P3b from typical auditory and visual stimuli. *Clinical Neurophysiology*, *110*(1), 24–30. https://doi.org/10.1016/S0168-5597(98)00033-1

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. *Journal of Cognitive Neuroscience*, *21*(3), 550–566. https://doi.org/10.1162/jocn.2009.21036

Fischler, I., Jin, Y. S., Boaz, T. L., Perry, N. W., & Childers, D. G. (1987). Brain potentials related to seeing one's own name. *Brain and Language*, *30*(2), 245–262. https://doi.org/10.1016/0093-934X(87)90101-5

Friston, K. J., Rotshtein, P., Geng, J. J., Sterzer, P., & Henson, R. N. (2006). A critique of functional localisers. *NeuroImage*, *30*(4), 1077–1087. https://doi.org/10.1016/j.neuroimage.2005.08.012

Holeckova, I., Fischer, C., Giard, M. H., Delpuech, C., & Morlet, D. (2006). Brain responses to a subject's own name uttered by a familiar voice. *Brain Research*, *1082*(1), 142–152. https://doi.org/10.1016/j.brainres.2006.01.089

Hommel, B., & Akyürek, E. G. (2005). Lag-1 sparing in the attentional blink: Benefits and costs of integrating two events into a single episode. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *58*(8), 1415–1433. https://doi.org/10.1080/02724980443000647

Jones, W., Pincham, H., Gootjes-Dreesbach, E. L., & Bowman, H. (2020). Fleeting Perceptual Experience and the Possibility of Recalling Without Seeing. *Scientific Reports*, *10*(1). https://doi.org/10.1038/s41598-020-64843-2

Lawrence, D. H. (1971). Two studies of visual search for word targets with controlled rates of presentation. *Perception & Psychophysics*, *10*(2), 85–89. https://doi.org/10.3758/BF03214320

Luck, S. (2005). *An introduction to the event-related potential technique*. The MIT Press.

Mack, A., & Rock, I. (1998). Inattention: Faces and Other "Meaningful" Stimuli. In *Inattentional Blindness*. The MIT Press. https://doi.org/10.7551/mitpress/3707.003.0009

Martens, S., Munneke, J., Smid, H., & Johnson, A. (2006). Quick minds don't blink: Electrophysiological correlates of individual differences in attentional selection. *Journal of Cognitive Neuroscience*, *18*(9), 1423–1438. https://doi.org/10.1162/jocn.2006.18.9.1423

Olivers, C. N. L., Van Der Stigchel, S., & Hulleman, J. (2007). Spreading the sparing: Against a limited-capacity account of the attentional blink. *Psychological Research*, *71*(2), 126–139. https://doi.org/10.1007/s00426-005-0029-z

Pickering, E. C., & Schweinberger, S. R. (2003). N200, N250r, and N400 Event-Related Brain Potentials Reveal Three Loci of Repetition Priming for Familiar Names. *Journal of Experimental Psychology: Learning Memory and Cognition*, *29*(6), 1298–1311. https://doi.org/10.1037/0278-7393.29.6.1298

Pincham, H. L., 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. https://doi.org/10.1016/j.cortex.2016.04.007

Polich, J., & Kok, A. (1995). Cognitive and biological determinants of P300: an integrative review. *Biological Psychology*, *41*(2), 103–146. https://doi.org/10.1016/0301-0511(95)05130-9

Proverbio, A. M., Mariani, S., Zani, A., & Adorni, R. (2009). How are "Barack Obama" and

"President Elect" differentially stored in the brain? An ERP investigation on the processing of proper and common noun pairs. *PLoS ONE*, *4*(9), 7126. https://doi.org/10.1371/journal.pone.0007126

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary Suppression of Visual Processing in an RSVP Task: An Attentional Blink? *Journal of Experimental Psychology: Human Perception and Performance*, *18*(3), 849–860. https://doi.org/10.1037/0096-1523.18.3.849

Rosenfeld, J. P. (2005). "Brain Fingerprinting:" A Critical Analysis. *The Scientific Review of Mental Health Practice*, *4*(1), 1–35. http://groups.psych.northwestern.edu/rosenfeld/NewFiles/BFcritiquerevsub3-6.pdf

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

Rosenzweig, G., & Bonneh, Y. S. (2020). Concealed information revealed by involuntary eye movements on the fringe of awareness in a mock terror experiment. *Scientific Reports*, 10(1). https://doi.org/10.1038/s41598-020-71487-9

Schweinberger, S. R., Pickering, E. C., Burton, A. M., & Kaufmann, J. M. (2002). Human brain potential correlates of repetition priming in face and name recognition. *Neuropsychologia*, *40*(12), 2057–2073. https://doi.org/10.1016/S0028-3932(02)00050-7

Simione, L., Akyürek, E. G., Vastola, V., Raffone, A., & Bowman, H. (2017). Illusions of integration are subjectively impenetrable: Phenomenological experience of Lag 1 percepts during dual-target RSVP. *Consciousness and Cognition*, *51*, 181–192.

https://doi.org/10.1016/j.concog.2017.03.004

Stenberg, G., Hellman, J., Johansson, M., & Rosén, I. (2009). Familiarity or conceptual priming: Event-related potentials in name recognition. *Journal of Cognitive Neuroscience*, *21*(3), 447–460. https://doi.org/10.1162/jocn.2009.21045

Tacikowski, P., Jednoróg, K., Marchewka, A., & Nowicka, A. (2011). How multiple repetitions influence the processing of self-, famous and unknown names and faces: An ERP study. *International Journal of Psychophysiology*, *79*(2), 219–230. https://doi.org/10.1016/j.ijpsycho.2010.10.010

Tacikowski, P., & Nowicka, A. (2010). Allocation of attention to self-name and self-face: An ERP study. *Biological Psychology*, *84*(2), 318–324. https://doi.org/10.1016/j.biopsycho.2010.03.009

Tateuchi, T., Itoh, K., & Nakada, T. (2012). Neural mechanisms underlying the orienting response to subject's own name: An event-related potential study. *Psychophysiology*, *49*(6), 786–791. https://doi.org/10.1111/j.1469-8986.2012.01363.x

Vogel, E. K., Luck, S. J., & Shapiro, K. L. (1998). Electrophysiological Evidence for a Postperceptual Locus of Suppression during the Attentional Blink. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(6), 1656–1674. https://doi.org/10.1037/0096-1523.24.6.1656

Wyble, B., Bowman, H., & Nieuwenstein, M. (2009). The Attentional Blink Provides Episodic Distinctiveness: Sparing at a Cost. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(3), 787–807. https://doi.org/10.1037/a0013902

Yang, H., Wang, F., Gu, N., Gao, X., & Zhao, G. (2013). The cognitive advantage for one's own name is not simply familiarity: An eye-tracking study. *Psychonomic Bulletin and Review, 20*(6), 1176–1180. https://doi.org/10.3758/s13423-013-0426-z

Zoumpoulaki, A., Alsufyani, A., & Bowman, H. (2015). Resampling the peak, some dos and don’ts. *Psychophysiology*, *52*(3), 444–448. https://doi.org/10.1111/psyp.12363

Appendix A

Discussion of the Programming Inconsistencies

Due to a programming error, Participants 2 and 3 were shown one different probe name pair and Participant 4 was shown two different probe name pairs to the other participants. We do not believe this error affected these participants' results. Both participants 2 and 3 scored 5/5 for probe recall and 0/5 for irrelevant recall, had high recognition scores for probe and low for the irrelevant, and had significant p-values, similar to the majority of participants that saw the original probe names. This suggests the error had no effect on their results. Participant 4 did not have a significant p-value but still recalled 4 of the 5 probes and none of the irrelevants, and scored high in the recognition test for probe (4.4) and low for irrelevant (1.6), showing that they still saw and recognised the probes but not the irrelevants. In comparison, participant 1 was presented with all of the original probe names and had a significant p-value but scored lower for recall (2/5) and recognition (3.8) of the probes, so it seems unlikely that the programming error would have had an effect and been the reason behind participant 4's non-significant p-value. Instead, noise in the ERP data is the more likely reason.

We also ran the ERP analyses while excluding the participants who saw different probe names in error and all results were the same (group level, recall and recognition) or better (10/11 significant at the individual level). This shows that the programming error did not have a substantive effect on the results or impact our conclusions.

Appendix B

Discussion of the Fz and Cz Results, and the P3a

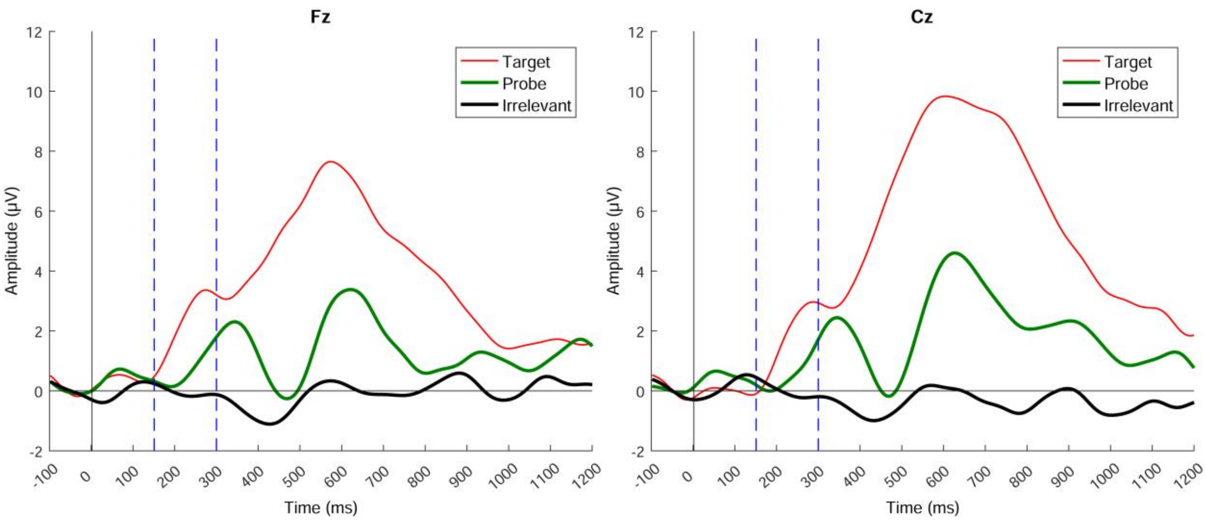

Figure 8. Grand Averages from the Fz and Cz Electrodes.

Group level analyses of Fz and Cz revealed no statistically significant difference between probe and irrelevant within a time window of 150 - 300ms (marked with blue dashed lines). Extra smoothing with a low pass filter of 10 Hz was used in this figure, purely for presentational purposes, and was not applied to the time-series analysed.

In previous Fringe-P3 studies (Bowman et al., 2013, 2014), significant P3a's (or novelty P3) were elicited at the Fz and Cz electrodes by own-name stimuli within a 150-300ms window, whereas in the present study, no significant difference was found between the probe and irrelevant conditions at Fz (P3a-Fz: $t(13) = 2.014$, $p = 0.065$, Cohen's $d = 0.766$) or Cz (P3a-Cz: $t(13) = 1.537$, $p = 0.148$, Cohen's $d = 0.608$), within the same P3a time window. Grand average ERPs from Fz and Cz are presented in figure 8. Several studies have shown that the P3a is often elicited in response to own-name stimuli (Fischler et al., 1987; Holeckova et al., 2006; Tateuchi et al., 2012). However, studies using other names such as Schweinberger et al. (2002) found no significant difference between familiar and unfamiliar names at frontal or central sites, and Tacikowski and Nowicka (2010) found that familiarity effects were more significant at parietal than at frontal or central sites. It can be clearly seen in our data (see figures 2 and 8) that P3 amplitude was highest at Pz. It is worth noting, however, that comparisons between the P3 waveform in the present study and Bowman et al. (2013, 2014) may be limited due to differences in the stimuli (this study used multi-item stimuli while Bowman et al used single item stimuli) and experimental tasks (this study had no explicit task relating to the famous names, while Bowman et al instructed participants to respond deceptively to their own names), which could have contributed to the differences in the waveforms.

Author Notes

Funding: One PhD student, KH, was supported financially by the University of Kent.

Declarations of interest: none.

Data Availability: The data that support the findings of this study are openly available in the Breakthrough Percepts of Famous Names repository on the Open Science Framework at https://osf.io/k6a5b/
DOI 10.17605/OSF.IO/K6A5B