Received: 11 November 2020 | Revised: 20 December 2020 | Accepted: 23 December 2020
DOI: 10.1111/ejn.15098

SHORT COMMUNICATION

EJN European Journal of Neuroscience FENS WILEY

Breakthrough percepts of online identity: Detecting recognition of email addresses on the fringe of awareness

Kathryn Harris[1] | **Claire Miller**[2] | **Brendan Jose**[2] | **Anthony Beech**[2] | **Jessica Woodhams**[2] | **Howard Bowman**[1,2]

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

Correspondence
Kathryn Harris, School of Computing, University of Kent, Canterbury, Kent, CT2 7NF, United Kingdom.
Email: kh468@kent.ac.uk

Funding information
We recognise funding support from the "Better Policing Collaborative Project" from the College of Policing, HEFCE, and Home Office to HB and JW, as well as an Experimental Psychology Society small grant to HB, AB, and JW. One PhD student, KH, was supported financially by the University of Kent.

Abstract
A key issue facing cybercrime investigations is connecting online identities to real-world identities. This paper shows that by combining the Fringe-P3 method with a concealed information test, we can detect a participant's familiarity with their own email address, thus connecting their real-world identity to their online one. Participants were shown Rapid Serial Visual Presentation (RSVP) streams of email addresses, some including their own email address (probe) or a target email address. Familiarity with the probe was accurately detected with significant results at the group level and for 7 of 11 participants at the individual level. These promising results demonstrate that the method can be successfully used to detect online identities. Factors that may affect how well an email address probe stands out in the RSVP streams are also discussed.

KEYWORDS
concealed information test, cybercrime, deception detection, Fringe-P3, Rapid Serial Visual Presentation

1 | INTRODUCTION

The Fringe-P3 method has been proposed as a tool for use in a range of forensics applications (Bowman et al., 2013). For example, it has been shown that a concealed information test (CIT) can be implemented with the Fringe-P3 approach (Bowman et al., 2013), which, importantly, is resistant to key countermeasures (Bowman et al., 2014). The CIT presents a salient probe stimulus infrequently amongst frequent non-salient stimuli. Only a "guilty" participant would recognise the probe and evoke a P3 event-related potential (ERP) component, thereby, revealing that they have knowledge of a stimulus that they may be attempting to conceal, while "innocent" participants would not recognise the probe and generate a P3 response.

There is a large body of research using CITs with EEG to detect familiarity with a range of stimuli (see Rosenfeld (2020) for a review) but these have not used RSVP. RSVP is a method of presenting stimuli so rapidly that they are on the fringe of awareness. This is a key part of the Fringe-P3 method as typically only salient stimuli breakthrough into conscious awareness (Bowman & Avilés, in preparation; Bowman et al., 2013), meaning that non-salient stimuli are not usually consciously perceived, thus rendering key countermeasures ineffective (Bowman et al., 2014). RSVP also allows for the experiment to be run quickly, which is important to the advisors that we have from law enforcement. The combined use of a CIT and RSVP without EEG has been shown to accurately detect familiarity with face stimuli using the attentional blink (Ganis & Patnaik, 2009), and a related method to the Fringe-P3 that presented face, name, and word stimuli on the fringe of awareness has been shown to detect concealed information based on involuntary eye movements

(Rosenzweig & Bonneh, 2019, 2020). Our initial demonstrations of the Fringe-P3 method showed that it could work with names (Bowman et al., 2013, 2014). Additionally, we have now shown that the method also works with famous face stimuli (Alsufyani et al., 2019).

An important application, though, of the Fringe-P3 method could be to associate real-world individuals with their online identities. Those policing the Internet face the problem that they are often in possession of an online identity associated with criminal activity, but they are unable to link that identity to a real-world individual. In response, this paper provides a proof of principle demonstration that the Fringe-P3 method could also be applied to link real-world individuals to their online identities by detecting their familiarity with that online identity.

We did this by showing that a participant's own email address (the probe) evokes a P3 component when presented in RSVP amongst unfamiliar email address distractors. This P3 response is contrasted with the absence of a P3 response to an unfamiliar email address (the irrelevant) that has the same presentation statistics as the probe.

More specifically, we demonstrate a very strong group level effect of probe (own email address) versus irrelevant (unfamiliar email address). We, then, show evidence of a strong effect at the level of individual participants using our standard permutation procedure (see Method section; Alsufyani et al., 2019; Bowman et al., 2013, 2014). In both analyses, we used our standard Aggregated (Grand or ERP) Average of Trials (see Method section; Bowman et al., 2020; Brooks et al., 2017) as a means to identify analysis windows without inflating false-positive rates.

Our conclusions focus on the effectiveness of the group level and individual level analyses and the promise of the Fringe-P3 method for application to the investigation of cybercrime.

2 | METHOD

2.1 | Participants

Fifteen students from the University of Birmingham took part in the experiment and gave their informed consent. Three participants were excluded due to a technical error with the recording system. One further participant was excluded for having less than half the trials remaining after trial rejection, leaving eleven participants for analysis. The eleven participants were aged 19–28 ($M = 23.27$, $SD = 2.9$), seven female and four male. All participants had normal or corrected-to-normal vision and no known neurological disorders. The STEM Research Ethics Committee of the University of Birmingham granted ethical approval for this study.

2.2 | Stimuli

All stimuli were email addresses using the University of Birmingham's email format. University of Birmingham email addresses are generated according to a set of rules, using three letters from the student's initials followed by three numbers (e.g., ABC123 or CXD456). If the student does not have three initials from their names, an X is added between their two initials. All stimuli generated for this experiment matched this format. There were three critical stimuli: probe, target, and irrelevant, presented in streams amongst randomly generated distractor email addresses. The target stimulus was an email address generated at random that participants were to look for within the RSVP streams, which they needed to pretend was their own. The probe stimulus was the participant's real-email address, which they were not warned would appear within the streams. The irrelevant stimulus was a randomly generated unfamiliar email address, which, unlike the target, was not shown to the participants before the RSVP streams. These critical stimuli were presented amongst unfamiliar distractor stimuli. A bank of 3,667 distractors was randomly generated, with the condition that they could not have more than two characters the same as any of the critical stimuli. Distractors for each stream were selected pseudorandomly from this bank, while ensuring that each distractor could not contain more than one character in the same place as the stimulus before it. Each distractor had the same small chance (0.38%) of appearing in a stream. Distractors and irrelevant stimuli are the same in terms of generation procedure and unfamiliarity, but each individual distractor is presented very infrequently, while the irrelevant stimuli are presented as frequently as the probe and target stimuli.

2.3 | Stimulus presentation

The presentation method was based on previous Fringe-P3 research by Bowman et al. (2013, 2014) and Alsufyani et al. (2019). The stimuli were presented in streams of 15 items. One critical stimulus appeared in each stream, placed randomly between the fifth and eighth items (inclusive) of each stream. Distractors were pseudo randomly chosen to fill the rest of the stream with the limitation that a distractor could not be shown immediately after any stimulus with which it shared more than one identical character. Before the 15 email address stimuli, there was a fixation stimulus ("XXXXXX") in the centre of the screen to focus the participant's attention on the presentation area, and at the end, there was a final stimulus ("------" or "======") which was chosen randomly. After the final stimulus, participants were asked whether they saw the target (responding yes or no) and which of the two possibilities the final stimulus was.

This final stimulus question was to ensure that participants paid attention to the full length of the stream.

Streams were presented in five blocks of 36 trials. Participants were given the opportunity to take a break after each block. In each block, the critical stimuli were randomly presented 12 times each with one critical stimulus per stream. Over the five blocks, the target, probe, and irrelevant were presented 60 times each, with 180 trials in the whole experiment.

The Stimulus Onset Asynchrony (SOA) for each participant was chosen using a staircase procedure to find the speed where the hit rate for the target was 75% and the correct rejection rate was 80% for each participant during practice trials. The SOAs used in the experiment were within a range of 100–250 ms.

2.4 | Recording apparatus

The data were recorded using the BioSemi ActiveTwo system and ActiView software. Electroencephalographic data were recorded from Fz, Cz, and Pz. Electrooculogram data were recorded from both eyes using two HEOG electrodes on the outer canthus of each eye and two VEOG electrodes, one above and one below the right eye. Linked mastoids were used as a reference. Impedances were kept below 10 kOhm. The data were digitised at 2,048 Hz.

2.5 | Analysis procedure

The data were analysed from the Pz channel using MATLAB 2016a and EEGLAB v13.6.5b. The data were resampled to 512 Hz and underwent a high-pass filter of 0.5 Hz and low-pass filter of 30 Hz, before being epoched into segments from −100 ms to 1,500 ms. Trial rejection was performed on the eye electrodes (with above 100 μV and below −100 μV as criteria), and on the Pz channel (with above 50 μV and below −50 μV as criteria). Baseline correction was then performed on −100 ms to 0 ms.

After trial rejection, the maximum number of trials remaining per participant for probe and irrelevant was 60 and the minimum remaining was 50 for probe (M = 56, SD = 3.46) and 43 for irrelevant (M = 54.91, SD = 5.43).

2.6 | AGAT analysis method

Only the probe and irrelevant data were used in the analyses. These were compared with each other to see if there was a statistically significant difference between them. The target trials were only included in the experiment to provide a task for the participants and the distractor stimuli were only included to fill the streams. The target and distractor stimuli provide no information on the recognition of a participant's own email address, and so were not used in any analyses.

2.6.1 | Group level

Statistical analysis was conducted according to the Aggregated Grand Average of Trials (AGAT) method (Bowman et al., 2020; Brooks et al., 2017). For this, the trials for the probe and irrelevant conditions were merged to provide one aggregated ERP per participant. A window 100 ms wide was placed on each participant's aggregated ERP, starting at 250 ms, and slid along by just under 1.95 ms each time, until the upper boundary was reached at 1,000 ms. The 250 ms start point was chosen to allow for individual differences around the start of the P3 response (typically around 300 ms). The mean amplitude of the aggregated ERP within each window position was calculated to find the window of interest (i.e., the window with the highest mean amplitude). This window of interest was then applied separately to the probe ERP and irrelevant ERP for that participant, and the mean amplitude within that window was calculated for each to then be analysed using a paired-samples t test.

Finding the window of interest from the aggregated ERP prevents any bias towards probe or irrelevant or inflation of the false-positive rate. The key step in this procedure is the merging of probe and irrelevant. Essentially, as verified in simulation and proof (Bowman et al., 2020), this merging means that, under the null, selecting this window of interest from the resulting aggregated ERP is mathematically orthogonal to the probe minus irrelevant contrast. Accordingly, under the null, there is no increased probability of the probe being different from irrelevant at the window selected on the aggregated ERP.

2.6.2 | Individual level

The individual level analysis also used the window of interest from the aggregated ERP and applied this to the individual probe and irrelevant ERPs to get the mean amplitudes from that window. The probe mean minus the irrelevant mean was calculated as the true observed difference. The individual probe and irrelevant trials were then permuted 1,000 times to create surrogate probe and irrelevant ERPs. The same window of interest as before was then applied to the surrogate probe and surrogate irrelevant ERPs, their means calculated, and the difference of the means calculated as surrogate differences. The p-value was the number of surrogate differences that were greater than or equal to the true difference, divided by the number of permutations.

Permutation was used for the individual participants' analysis, as the generation of surrogate ERPs have a better signal-to-noise ratio than if *t* tests are performed on the single trials, which contain more noise.

3 | RESULTS

3.1 | Group effects

A grand average for the probe and irrelevant conditions from the Pz electrode is presented in Figure 1. It can be seen from Figure 1 that there is a clear P3 response for the probe condition that is not present for the irrelevant condition. The result of the group-level AGAT analysis and paired-samples *t* test found a significant difference and a large effect size between the probe and irrelevant conditions at Pz, $t(10) = 4.240$, $p = .002$; $d = 1.824$.

Previous Fringe-P3 studies using own-name stimuli (Bowman et al., 2013, 2014) have found significant P3a's at Fz and Cz within a 150–300 ms window. No significant difference was found between the probe and irrelevant during this window for the current study at Fz ($t(10) = 0.149$, $p = .885$, $d = 0.051$) or Cz ($t(10) = 0.235$, $p = .819$, $d = 0.075$). It may be that the particularly exquisite salience of one's own name, which a University email address may not possess, drove the generation of the P3a in our previous work.

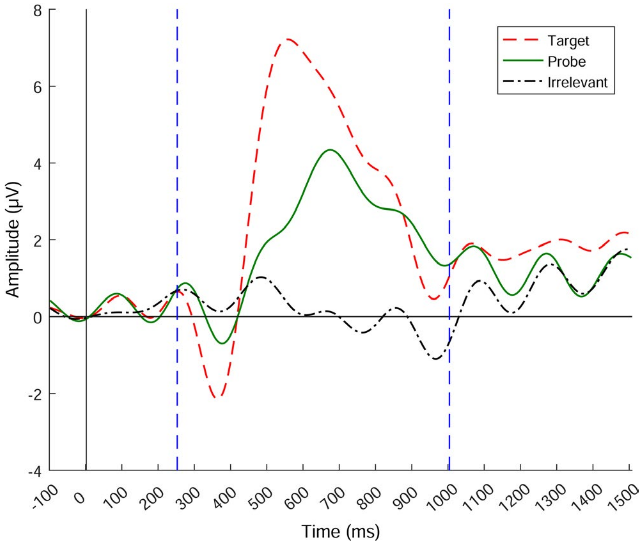

FIGURE 1 Grand Average for All Participants' Probe and Irrelevant from the Pz Electrode. Extra smoothing with a low-pass filter of 6 Hz was used in this figure, purely for presentational purposes, and was not applied to the time-series analysed. Vertical dashed lines represent the 250–1,000 ms area within which the AGAT slides.

3.2 | Individual participants

Figure 2 shows the ERPs for each individual participant and their window of interest from the AGAT analysis. Table 1 shows the *p*-values obtained from the AGAT permutation tests on the probe and irrelevant conditions from the Pz data. There were seven out of eleven participants (64%) with a significant difference between the probe and irrelevant at an alpha of 0.05, with four of these having *p*-values less than .001. The mean *p*-value was .115 and the median *p*-value was .034. The median is the more appropriate measure of central tendency, rather than the mean, due to the skewed distribution of *p*-values.

4 | DISCUSSION

This research presented an initial investigation into the use of the Fringe-P3 concealed information test for online identities, to link real-world identities to Internet identities (e.g., email addresses, Skype addresses, identifiers associated with online forums) connected to online crimes. The group level analysis found a significant difference between responses to the probe and irrelevant stimuli ($p = .002$), and a clear P3 response for the participants' own email addresses but not for the unfamiliar irrelevant email addresses. These results show that the method was successful in detecting when participants saw their own email address amongst non-familiar stimuli.

In addition to the group level analysis, a significant difference was found between the probe and irrelevant for seven out of eleven participants (64%) in the individual level analysis. Four of these participants had highly significant differences with *p*-values of .001 or smaller. These promising results highlight how accurately this method can work for an individual participant as well as at the group level. Both the group and individual level results demonstrate the strength and suitability of the Fringe-P3 method and AGAT analysis in the detection of people's online identities.

While these results are generally significant at the individual level, as well as at the group level, previous Fringe-P3 research using own names (Bowman et al., 2013, 2014) and faces (Alsufyani et al., 2019) have found even stronger results at the individual level. There are a few possible reasons why four participants did not have significant differences between their probe and irrelevant in this individual analysis. The following explanations warrant further research that may result in improvements to make the Fringe-P3 concealed information test more successful.

One potential explanation is that participants' university email addresses may not be as salient to them as expected, especially compared to own names (Bowman et al., 2013, 2014). Our own names have significant meaning to us, making them highly salient and stand out more

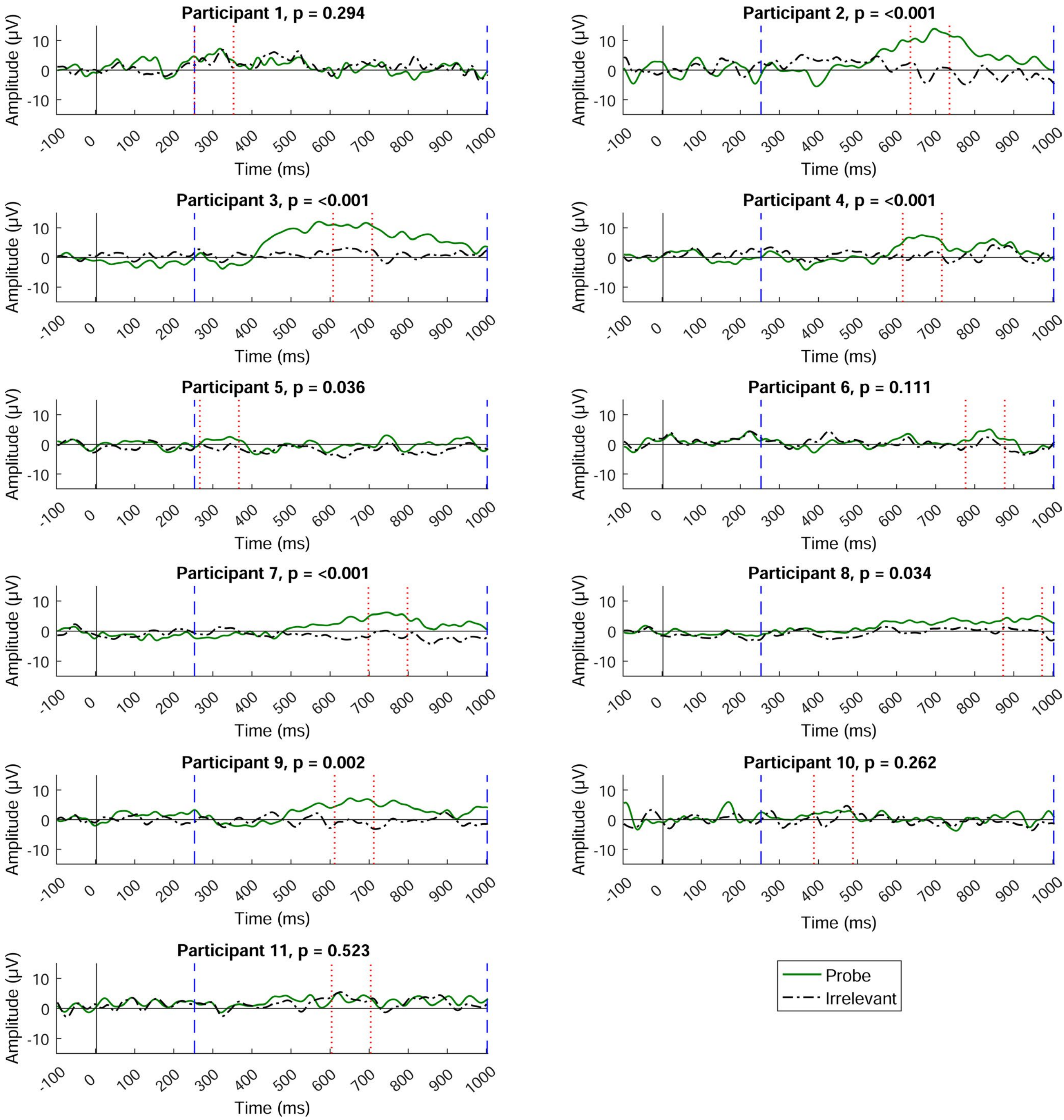

FIGURE 2 Individual Participants' ERPs from the Pz Electrode. Dashed vertical lines represent the 250–1,000 ms area within which we searched for the analysis window. Dotted vertical lines represent the window of interest found in the AGAT analysis for that participant.

easily amongst distractors, to the point of even escaping an attentional blink when presented shortly after another critical stimulus (Shapiro et al., 1997). In contrast, University of Birmingham email addresses are generated according to university rules, and not chosen by the participant, so they may not have the same level of meaning and salience to all participants and this may have contributed to why they did not stand out as significantly in the streams compared to a participant's own name.

There is also a potential perceptual issue with the email address stimuli, since they are generated following the same rules and are the same length, making them very homogenous. Additionally, due to the generation rules, several stimuli have an X in the middle, further adding to their homogeneity. When distractors are too similar to the target (e.g., same category or colour), the distractors can capture attention (Folk et al., 1992; Su et al., 2011). There needs to be heterogeneity in RSVP stimuli in order for the critical salient stimuli to

TABLE 1 Individual participants' *p*-values from the AGAT analysis

Participant	*p*-value
1	.294
2	<.001*
3	<.001*
4	<.001*
5	.036*
6	.111
7	<.001*
8	.034*
9	.002*
10	.262
11	.523
Mean *p* value	.115
Median *p* value	.034

*Significant at the .05 level.

stand out from the distractors. It is possible the homogeneity of the email address stimuli may have contributed to some participants not demonstrating a significant effect, compared to previous research that used more heterogeneous stimuli, such as names and faces (Alsufyani et al., 2019; Bowman et al., 2013, 2014).

Finally, there is the possibility that participants were (whether consciously or not) using a search strategy when looking for the target, such as looking for the first letter of the target email address and only processing the rest of the stimuli if the first letter matched. This could mean they did not read the probe if it did not start with the same letter as the target, which would prevent a P3 response. This fits with the glance-look model of cognitive control (Su et al., 2011), which argues that stimuli in RSVP streams are first "glanced" at to process the broad category or meaning of the stimulus (in our example, if it begins with a specific letter). If a stimulus matches the target category then it receives a deeper "look," where specific meaning and detail are analysed (and the rest of that email address would be read). Additionally, the glance-look theory posits that distractor stimuli that are in the same or a semantically related category as the target can capture attention, receive a deeper "look," and initiate an attentional blink, causing participants to miss a target or probe stimulus presented shortly after (Folk et al., 1992; Su et al., 2011). In the case of email addresses, this could mean that distractors that began with the same letter as the target could have captured attention, reducing the chance of the participant seeing the probe and invoking a P3 if it appeared shortly after. A future experiment should attempt to prevent this search strategy by finding a task that forces participants to read the whole email address.

Research into these potential issues outlined above could lead to improvements in both the stimuli and procedures used for future online identity deception experiments and hence could potentially result in even stronger individual participant level responses. Despite these possible issues, the current study found significant results at both the group and individual level, thus providing the intended proof of principle demonstration that this method can successfully detect the recognition of a participant's own email address. This is a vital step towards linking online identities to real-world identities and aiding those fighting cybercrime. This can be performed in forensic situations where the police know an online identity is associated with a crime, and they suspect that a particular real-world person is the user of that online identity. The Fringe-P3 method can then be used with that real-world person to demonstrate their familiarity with that online identity and link them to it.

CONFLICT OF INTEREST

All authors have declared no conflict of interest.

AUTHOR CONTRIBUTIONS

Kathryn Harris processed the data, ran the statistical analyses, and wrote the manuscript. Claire Miller coded the experiment and collected the data. Jessica Woodhams, Anthony Beech, Brendan Jose, and Howard Bowman contributed to formulating the idea for the experiment and to its design. Howard Bowman also contributed to the analysis methods used. All authors reviewed the manuscript.

Peer Review

The peer review history for this article is available at https://publons.com/publon/10.1111/ejn.15098.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in the "Breakthrough Percepts of Online Identity" repository on the Open Science Framework at https://osf.io/skemh/ https://doi.org/10.17605/OSF.IO/SKEMH

ORCID

Kathryn Harris https://orcid.org/0000-0003-3576-5970
Claire Miller https://orcid.org/0000-0001-5040-1441
Anthony Beech https://orcid.org/0000-0002-6681-802X
Jessica Woodhams https://orcid.org/0000-0002-9674-6653
Howard Bowman https://orcid.org/0000-0003-4736-1869

REFERENCES

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

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*, *16*(11), e1008286. https://doi.org/10.1371/journal.pcbi.1008286

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

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

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary Covert Orienting Is Contingent on Attentional Control Settings. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(4), 1030–1044. https://doi.org/10.1037/0096-1523.18.4.1030

Ganis, G., & Patnaik, P. (2009). Detecting concealed knowledge using a novel attentional blink paradigm. *Applied Psychophysiology Biofeedback*, *34*(3), 189–196. https://doi.org/10.1007/s10484-009-9094-1

Rosenfeld J. P. (2020). P300 in detecting concealed information and deception: A review. *Psychophysiology*, *57*(7), 1–12. https://doi.org/10.1111/psyp.13362

Rosenzweig, G., & Bonneh, Y. S. (2019). Familiarity revealed by involuntary eye movements on the fringe of awareness. *Scientific Reports*, *9*(1), 1–12. https://doi.org/10.1038/s41598-019-39889-6

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), 1–15. https://doi.org/10.1038/s41598-020-71487-9

Shapiro, K. L., Caldwell, J., & Sorensen, R. E. (1997). Personal Names and the Attentional Blink: A Visual, "Cocktail Party" Effect. *Journal of Experimental Psychology: Human Perception and Performance*, *23*(2), 504–514. https://doi.org/10.1037/0096-1523.23.2.504

Su, L., Bowman, H., & Barnard, P. (2011). Glancing and then looking: On the role of body, affect, and meaning in cognitive control. *Frontiers in Psychology*, *2*(DEC), 1–48. https://doi.org/10.3389/fpsyg.2011.00348

How to cite this article: Harris K, Miller C, Jose B, Beech A, Woodhams J, Bowman H. Breakthrough percepts of online identity: Detecting recognition of email addresses on the fringe of awareness. *Eur J Neurosci*. 2021;53:895–901. https://doi.org/10.1111/ejn.15098