The Plymouth Student Scientist - Volume 08 - 2015 The Plymouth Student Scientist - Volume 8, No. 2 - 2015

2015

Why is eyewitness testimony so inaccurate? An investigation into event-related potentials and the recognition of familiar and unfamiliar faces

Evans, J.

Evans, J. (2015) 'Why is eyewitness testimony so inaccurate? An investigation into event-related potentials and the recognition of familiar and unfamiliar faces', The Plymouth Student Scientist, 8(2), p. 133-148.
http://hdl.handle.net/10026.1/14100

The Plymouth Student Scientist
University of Plymouth

*The Plymouth Student Scientist, 2015, **8**, (2), 133-148*

Why is eyewitness testimony so inaccurate? An investigation into event-related potentials and the recognition of familiar and unfamiliar faces

Jack Evans

Project Advisor: Giorgio Ganis, School of Psychology, Plymouth University, Drake Circus, Plymouth, PL4 8AA

Abstract

Eyewitness-testimony is notoriously inaccurate and unreliable. In addition, a body of research has indicated that unfamiliar faces are poorly recognised in comparison to famous or familiar faces. Under the assumption that poor unfamiliar face recognition plays a role in the inaccuracy of eyewitness testimony, the present study sought to use the P300 ERP to provide evidence that unfamiliar face recognition is a factor in the inaccuracy of eyewitness testimony. After showing videos of either busy public places, mock-crime videos, or a film clip, this experiment recorded EEG data from 12 healthy undergraduates in their 20s attending Plymouth University whilst they saw photographs of famous and unfamiliar faces appear on a screen. Included in the photographs were faces they had seen in the videos, and after this, participants had to make a line-up identification of the faces they had seen in the videos. The results indicated that poor unfamiliar face recognition is a factor in the inaccuracy of eyewitness testimony, based on the P300 and line-up responses of the participants to unfamiliar faces. The results also highlighted possible problems with the methodology used in this study. Future research should consider improving upon the methodology used in this study to clarify and provide further support for the argument presented in this paper.

Introduction

On eyewitness testimony

Eyewitness testimonies are an important source of information and are one of the most commonly used forms of evidence in criminal investigations. Remaining highly regarded amongst law professionals (Brigham and WolfsKeil, 1983), and often substantially impacting upon jurors decisions (Devine, Clayton, Dunford, Seying, and Pryce, 2001), there is no mistake that eyewitness testimony can be crucial in determining innocence or guilt in the courtroom. Opinions amongst police officers reach agreement that eyewitness testimony often provides police with a central lead in their investigations (Kebbell and Milne, 1998) and can tie a suspect directly to a crime (Wells & Loftus, 2003). It is clear that eyewitness testimony is an incredibly useful and important source of information in criminal cases. Despite this, there is controversy surrounding the use of eyewitness accounts as evidence in court.

Eyewitness testimony is notoriously unreliable and studies have shown that the rates of an eyewitness misidentifying a guilty suspect can be worryingly high (Buckhout, 1980; Wells, 1993). In addition to this, statistics indicate that eyewitness testimony was a leading cause in over 75% of wrongful convictions (Scheck, Neufield and Dwyer, 2000; The Justice Project, 2007). This poses a problem; poor performance from an eyewitness could mean time in jail for an innocent person, or freedom for a murderer. There's no doubt that eyewitness testimony is and will continue to be a vital source of evidence in criminal proceedings, but why is it so inaccurate, and can its accuracy be improved?

Event-related potentials

The ability of an eyewitness to successfully identify a culprit essentially relies on their memory, their ability to remember or recognise who it was that they saw commit the crime, and to do this well enough to produce a match. Herein lies the problem; eyewitnesses identify someone as the culprit even without truly recognising them.

One way to improve recognition, or rather to detect whether an eyewitness recognises a culprit, may be through the analysis of Event-Related Potential (ERP) components. ERP components are voltage changes occurring in the brain in response to cognitive events or perceived stimuli (Blackwood and Muir, 1990; Luck, 2005). The literature identifies a specific ERP component called the P300, which is elicited in response to rare and meaningful stimuli (Johnson, 1993). It is possible, therefore, that the P300 component can be used to identify whether or not a participant recognises something by the amount of meaning it may have to them.

Self-relevant and autobiographical information

Previous research has shown how the presentation of self-relevant stimuli can elicit a stronger P300 ERP response than non-relevant stimuli. Berlad and Pratt (1995) measured ERP responses whilst presenting participants with their name and two irrelevant words. The participants name and one of the irrelevant words were presented infrequently, whilst the other irrelevant word was presented more often. They found that the participants' name elicited a larger P300 ERP component than both the infrequent irrelevant word and the frequently-presented irrelevant word. This suggests that stimuli with meaning to the participant can be neurologically distinguished from other rare and infrequent information, and can be identified by a larger P300.

Gray, Ambady, Lowenthal, and Deldin (2004) provided support for this through their investigation into whether self-relevant stimuli are still recognised when attention is directed elsewhere. It was found that despite attentional resources being directed to another type of stimulus, information that was relevant to the participant (such as their name or home-town) elicited a large P300 that was three times larger than irrelevant stimuli. Such a finding provides further evidence that the P300 can be used to distinguish meaningful stimuli from those that do not evoke meaning at all, and this difference can identify information that participants recognise. This finding for autobiographical information implies that the P300 may elicit similar amplitude for other information that has meaning. For example, the recognition of information you are trying to deny knowledge of, or the recognition of objects related to a crime.

Concealed information tests

This paradigm has been used in forensic contexts of concealed knowledge. Several studies have had participants take part in a simulated mock-crime and then try to deny knowledge of crime-related information; or to deny knowledge of self-relevant/autobiographical information. The basic principle behind these P300 concealed information tests is that information relevant to the individual will be meaningful to them; or that people who took part in a mock-crime will have knowledge related to the crime (such as the date the crime was committed, the weapon used or items stolen), and thus elicit a P300 when presented with this information.

Rosenfeld, Cantwell, Nasman, Wojdac, Ivanov, and Mazzeri (1988) asked participants to steal one of nine items from a box. During the recording of the ERPs, the names of objects would appear on screen; one of which being the ‘stolen’ object. A control group who did not steal anything also took part in the task. On analysis of the ERPs, it was found that only the guilty participants elicited a P300 to the ‘stolen’ item when it appeared on screen. This illustrated how the P300 can be used to detect deception; a larger P300 elicited to stolen stimuli indicates recognition of the stolen object, and thus this can be used to distinguish guilty participants from innocents.

Support for this was given by Meijer, Smulders, and Wolf (2009) through their investigation into concealed face recognition and mere-recognition effects. Using pictures of family and friends whilst ERPs were being recorded, the authors asked participants in one condition to deny recognition to some familiar faces. In another condition participants were required only to pay attention to a dot on the cheek of the familiar face, and indicate whether the dot was on the left or right cheek. Their results indicated that, despite any denial of recognition, faces that were familiar to the participant could be identified via the enlargement of the P300.

Furthermore, the results also indicated that when participants were instructed to respond to an irrelevant dot and not pay attention to the face, a P300 was still evident for these familiar faces in comparison to unfamiliar control faces. The results of this study give further evidence to how the P300 can be used to detect recognition of meaningful information. Furthermore, not only do the results support the use of pictures of faces in a P300 setting; but also that familiar faces such as family or friends can elicit a P300, based off recognition alone, without any instructional guidance. Such a result therefore indicates that the P300 can possibly be used to detect facial recognition; for example, in the context of eyewitness-testimony.

Eyewitness recognition and mock crime scenarios

Lefebvre, Marchand, Smith and Connolly (2007) tested if the P300 could be used as a measure of eyewitness-testimony identification accuracy across 3 time-delay conditions (only the no-delay condition is relevant and therefore discussed here). Participants were shown various mock-crime videos in which a culprit entered a room and stole a laptop, with brief opportunities (two-three seconds) to see the culprits face from the front. Following this, whilst ERPs were recorded participants were presented with repetitions of anonymous faces and the culprits face from the video, and required to press a button if they saw the culprit or an anonymous face. Participants were also required to provide confidence judgements of who they chose as the culprit. The authors found that when button presses were correct, and when confidence judgements were high, the culprit elicited a significantly larger P300 than irrelevant faces.

The results also showed that the majority of eyewitness identifications were correct; however 21% of P300 identifications could not correctly identify the culprit. This illustrates that ERPs can indeed be used to index the accuracy of eyewitness testimony, and also confirms the point outlined in the introduction; eyewitness testimony can be quite unreliable.

Previous research by the author

Based on the research by Lefebvre et al (2007), the author of the present study conducted a similar experiment using ERPs in an eyewitness testimony context (Evans and Woolcock, Unpublished). Participants watched a mock-crime video of a Caucasian male entering a room and looking around before stealing a laptop, money from a purse, and a touch screen tablet. Following this, whilst ERP data was recorded participants were presented with a series of photographs of anonymous male faces, including a photograph of the culprit from the mock-crime video. The aim of the study was to find a larger positive P300 deflection when the photograph of the culprit was presented, compared with the anonymous faces. However, in contrast to the study by Lefebvre, the results revealed no significant difference between the P300 for the culprit and the anonymous faces ($p > .05$).

It was proposed the reason for such a result is due to the poor ability of humans to recognise unfamiliar faces. Previous studies have shown that unfamiliar faces have a notoriously poor recognition rate in comparison to famous/familiar faces (Ellis, Shepard and Davies, 1979; Hancock, Bruce, and Burton, 2000; Megreya and Burton, 2006). The author proposed that this could be a factor as to why participants did not recognise the culprit in the video. It is possible that this unfamiliar face effect may be why Lefebvre et al (2007) found some incorrect identifications in their data. Furthermore, it is possible that this effect may be a factor in why eyewitness testimony is so inaccurate.

The present study

The research above has described how eyewitness testimony is often inaccurate and unreliable. It has been seen how the P300 ERP component can be used to identify rare, meaningful information, and personally-familiar faces. In an eyewitness testimony context, studies have also shown how the P300 can be used as index for detecting if participants recognise culprits seen in mock-crime videos.

Based on past research, the author of this paper postulates that inaccurate eyewitness testimonies may be due to humans' poor recognition of unfamiliar faces,

thus resulting in poor eyewitness testimony performance and a lower amplitude P300 ERP. Therefore, the present study attempts to investigate and provide evidence for this assumption as a possible reason as to why eyewitness testimony can be so inaccurate.

This will be achieved by comparing the P300 amplitude elicited by an unfamiliar culprit with anonymous irrelevant faces after watching a control video in which the culprit is not present, and then after watching a mock-crime video in which the culprit is present. To gain concrete evidence of recognition in an eyewitness setting, participants will also be required to identify the culprit from a line-up after seeing the mock crime video. To illustrate a contrast of how well people recognise known faces, the author also tests famous faces under the same methodology; comparing the P300 amplitude of a specific famous actor with other famous faces after watching a control video, and then after watching a film-clip in which the actor is present. A line-up consisting of famous faces will also be administered.

There are several hypotheses for this study. Based on previous research, it is hypothesized that the unfamiliar culprit will not show a larger P300 than anonymous irrelevant faces after either the control or mock-crime video. The famous actor will elicit a larger P300 than other famous faces only after participants see the film-clip in which he is present. The famous actor will elicit a larger P300 than the unfamiliar culprit after participants watch the control video and after watching videos in which the culprit and actor were present. Finally, participants will respond more accurately to the line-up task for famous faces than they will to the task for unfamiliar faces.

Method and materials

Participants

A total of 12 healthy undergraduates, recruited from Plymouth University, took part in the experiment for course credit and were used in the final analysis (8 females and 4 males, mean age = 20 years, SD = 2.4). A demographic questionnaire administered before the study confirmed that all participants had normal or corrected vision and no history of neurological or psychiatric disease.

Stimuli

A total of 4 videos were used in the experiment, videos A, B, C and D. In the control conditions, the videos were approximately 50 seconds long and depicted one of two scenes; either (A) a camera documenting a large queue of people waiting at a bus stop, or (B) people walking through a London street. The other 2 videos were used in the experimental conditions. In the unfamiliar face condition, the video (C) depicted a 50 second long non-violent mock crime, in which a Caucasian male would enter a room and look around before stealing a laptop, a tablet and money from a purse, and then leave the room. The culprit's face was fully visible from the frontal view for a total of roughly 3 seconds. In the famous-face condition, the video (D) was a 39 second long clip taken from a film, in which a camera would pan around an office, and follow a famous actor's narration and monologue as he walked through the room. The actors face was fully visible from the front for roughly 3 seconds in total. The sound had been removed from all videos, so that the participant could concentrate on the visual scene.

Five greyscale pictures were used in each of the famous and unfamiliar conditions. In the unfamiliar conditions, the pictures consisted of 4 irrelevant photographs of anonymous male faces and a screenshot of the culprits face from the mock-crime

video (video C). The irrelevant photographs matched the culprit in terms of age, hair colour, facial hair and skin tone. In the famous conditions, the photographs consisted of 4 irrelevant photographs of famous male faces, and a screenshot of the actors face from the video clip (video D). The irrelevant photographs in both conditions were edited so that they resembled screenshots, and had a similar brightness, scale, and angle to the photograph of the culprit or actors face. All faces were presented frontally, but facing off at a slight angle to resemble the screenshots from the videos. All pictures were of an oval shape, approximately 9 centimetres lengthways and 6.5 centimetres widthways, and were presented on a black background.

Two line-up screens were used in this study that corresponded to the two experimental conditions. The line-up in the unfamiliar condition (A) consisted of the 4 unfamiliar filler photographs and a photograph of the culprit from video C. The line-up in the famous condition (B) consisted of the 4 famous filler photographs and a photograph of the actor from video D. Photographs in the line-ups were shown simultaneously and formatted to be an equal distance from each-other, and from the edges of the screen. All stimuli were presented on a 22 inch monitor approximately 150cm away from the participant, at eye level.

Experimental design

The experiment used a Within-Subjects design in which there were 4 conditions. (1) Unfamiliar face control, (2) unfamiliar face experimental, (3) famous face control, and (4) famous face experimental. In each condition there was a video phase; in which the participant would watch a video, and a test phase; during which the EEG data was recorded whilst the participant was presented a series of pictures. Each participant completed all four conditions, one after the other. The order in which participants completed the conditions was counterbalanced, so that some completed the famous control and experimental conditions first, whilst others completed the unfamiliar control and experimental first. The control conditions were always completed before the experimental condition. As an example, a participant may complete the unfamiliar control condition, then the unfamiliar experimental condition, followed by the famous control condition, and then the famous experimental condition.

Procedure

As a general overview: after the electrode setup was completed (detailed below), participants were seated in a comfortable chair and instructed to pay attention to a video. Following this, the participant began the test phase. During the test phase, the participant would see approximately 201 trials appear on the screen. A trial consisted of the presentation of a fixation cross, which remained on screen for approximately 650ms, followed by a photograph of a face which would remain onscreen for 800ms. The trials appeared sequentially and featured one of 5 photographs; the specific stimuli used in each condition is detailed below. The participant's task during the test phase was to respond via a button box when the same photograph of a face appeared twice in a row, on two consecutive trials; this was to ensure that participants paid attention to the stimuli. The pictures were presented pseudo-randomly (due to some intended repetitions), and each picture appeared roughly 35 times each over the course of the test phase.

In the unfamiliar face control condition, participants watched video A. In the unfamiliar face experimental condition, participants watched video C. During the test phase, the photographs presented in both the control and experimental conditions

included the faces of the 4 unfamiliar 'irrelevants' and the photograph of the culprit's face from video C, the mock crime video (see *Stimuli*).

In the famous face control condition, participants watched video B. In the famous face experimental condition, participants watched video D. During the test phase for these conditions, the photographs presented included the faces of the 4 famous irrelevants and the photograph of the actors face from video D, the film clip (see *Stimuli*).

A line-up task was administered to participants upon completion of each experimental condition. After completing the unfamiliar conditions, participants were shown line-up A, after completing the famous conditions, participants were shown line-up B. Participants were asked to look at each photograph in the line-up carefully, and to tell the experimenter which photograph depicted the person they saw in the corresponding video. They were told that the person may or may not be in the line-up, and were asked to tell the experimenter if they thought the person was not present, or if they did not know who they saw in the corresponding video.

Electrophysiological data acquisition

EEG was sampled at 250 Hz from Ag/AgCl electrodes (gain = 20,000, bandpass filtering = .01 to 100 Hz). EEG data were recorded from 32 electrodes on a cap which was placed on the participants' head, in accordance with the 10-20 electrode system. The electrode scalp sites were: Fp1, Fp2, AF3, AF4, F7, F3, Fz, F4, F8, T7, C3, Cz, C4, T8, CP5, CP1, CP2, CP6, P7, P3, Pz, P4, P8, PO3, PO4, O1, Oz, O2. These were referenced to the average of the mastoids. Facial electrodes were placed below the left eye to monitor eye blinks, lateral to the right and left eyes to monitor horizontal eye movements, and at the right and left mastoids to obtain the reference. Stimuli was presented using PresentationFront software, data was recorded (with a response box) using Biosemi Actiview, and analyses was conducted with the EEGLAB plugin for MATLAB, on laptop computers running Windows 7 and 8.

ERP analysis

ERPs were averaged off-line for an epoch of 1000ms, including a 200ms pre-stimulus baseline. Trials contaminated by blinks, eye movements, muscle activity, electrode drift, or amplifier blocking were rejected off-line. Following artifact rejection, a mean of 85.9% of artifact-free trials were kept for the analysis of the unfamiliar control condition, 85.1% for use in the unfamiliar experimental condition, 82.6% in the famous control condition, and 83.3% for use in the famous experimental condition. The ERP data from each participant was averaged together to form a grand average waveform for each of the four conditions. Visual inspection of these grand averages showed a centro-parietal positive deflection between 300-600ms. The three central and three parietal electrode sites with the highest P300 deflection were chosen for statistical analysis (C3, P3, Pz, P4, C4, Cz). A series of two-way repeated measures Analysis of Variance (RM ANOVA) were conducted on the mean amplitude of the averaged ERP data, between 300-600ms at these electrode sites, to assess the effects of the manipulations in the famous and unfamiliar conditions on the P300 component. The factors used in each ANOVA are detailed in their appropriate sections in the results section.

Results

The present study

The present study aimed to investigate possible reasons for the inaccuracy of eyewitness testimony. It was thought that inaccurate eyewitness testimonies were due to a lack of recognition of unfamiliar faces, as these are notoriously hard to recognise in comparison to known faces. Through the analysis of ERPs, this study tested the neurological basis of participants' ability to recognise specific unfamiliar and famous faces amongst irrelevant faces. They were to do this after watching either control videos of public places, or a mock-crime video or film clip in which an unfamiliar and famous face was present, respectively. In addition, a line-up task required participants to identify the unfamiliar face and the famous face that was shown to them in the videos.

There were several hypotheses for this study. It was predicted by the author that the unfamiliar culprit will not show a larger P300 than anonymous irrelevant faces after either the control or mock-crime video. In addition, a famous actor will elicit a larger P300 than other famous faces only after seeing the film-clip in which he is present. The author also hypothesized that the famous actor will elicit a larger P300 than the unfamiliar culprit after watching the control video and after watching videos in which the culprit and actor were present. Finally, participants will respond more accurately to the line-up task for famous faces than they will to the task for unfamiliar faces.

Grand average waveforms

Figures 1-4 depict the Grand Average waveforms for the unfamiliar control, unfamiliar experimental, famous control, and famous experimental conditions, respectively, at each of the six electrode sites (C3, P3, Pz, P4, C4 and Cz) for the culprit or actors face compared with the average irrelevant faces and the target. Figure 5 and 6 show the Grand Average waveforms of the unfamiliar culprit, the unfamiliar irrelevants, the famous actor, and the famous irrelevants, on the control condition and the experimental condition respectively. These plots allow for easier comparison of effects.

Unfamiliar conditions ANOVAs

As predicted for the unfamiliar conditions, a RM ANOVA comparing condition (unfamiliar culprit control, unfamiliar culprit experimental) and electrode site (C3, P3, Pz, P4, C4, Cz) as factors, revealed that there was no significant difference in P300 amplitude elicited by the culprit, between the unfamiliar control and the unfamiliar experimental condition, F (1, 11) = .40, $p > .05$.

In addition, a RM ANOVA conducted on the unfamiliar experimental condition using stimuli type (culprit, anonymous irrelevants), and electrode site as factors illustrated that participants showed no significant difference in P300 amplitude between the unfamiliar culprit and anonymous irrelevant faces, F (1, 11) = .002, $p > .05$. In line with the author's predictions, this indicated that participants were no better at recognising the culprit after seeing a mock-crime than beforehand, and that they recognised the culprit no better than anonymous irrelevant faces.

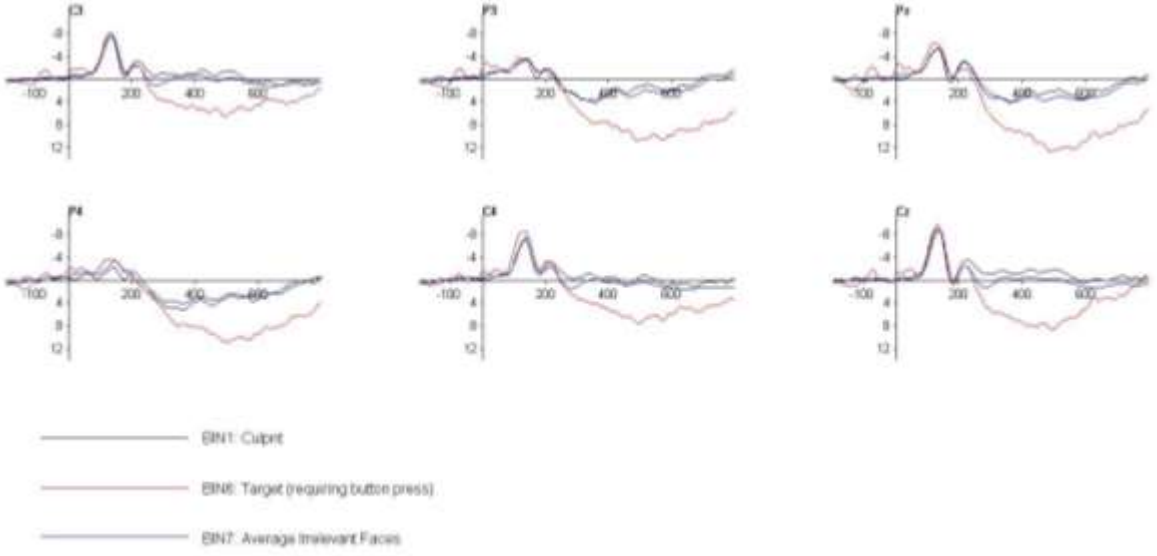

Figure 1: Unfamiliar control condition depicting grand average waveforms for the unfamiliar culprit, the target requiring a button press, and the average unfamiliar irrelevant faces

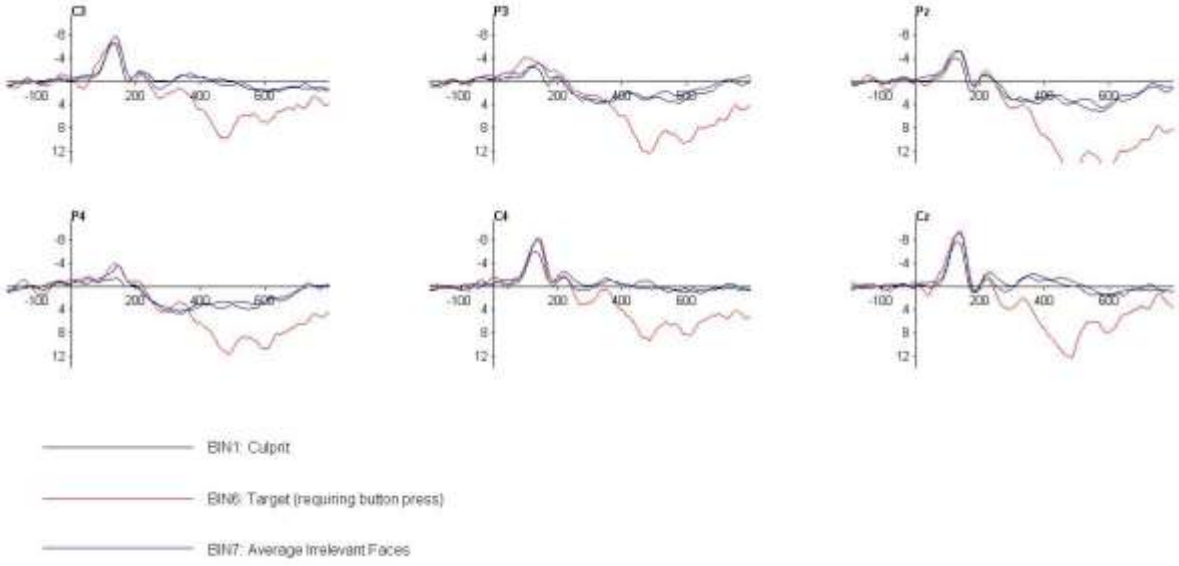

Figure 2: Unfamiliar experimental condition depicting grand average waveforms for the unfamiliar culprit, the target requiring a button press, and the average unfamiliar irrelevant faces

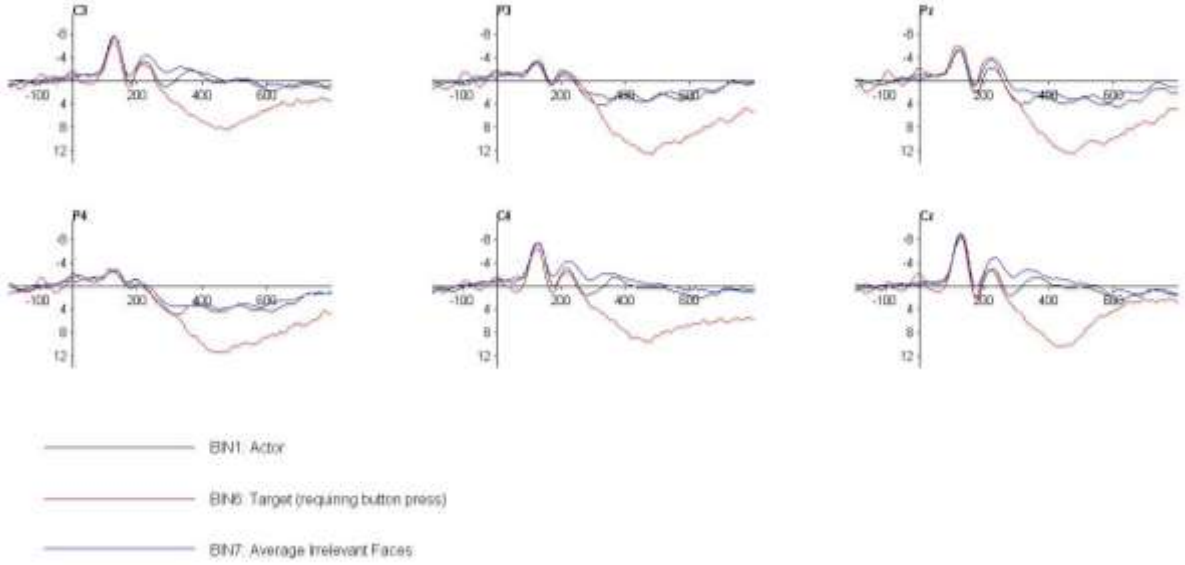

Figure 3: Famous control condition depicting grand average waveforms for the famous actor, the target requiring a button press, and the average famous irrelevant faces

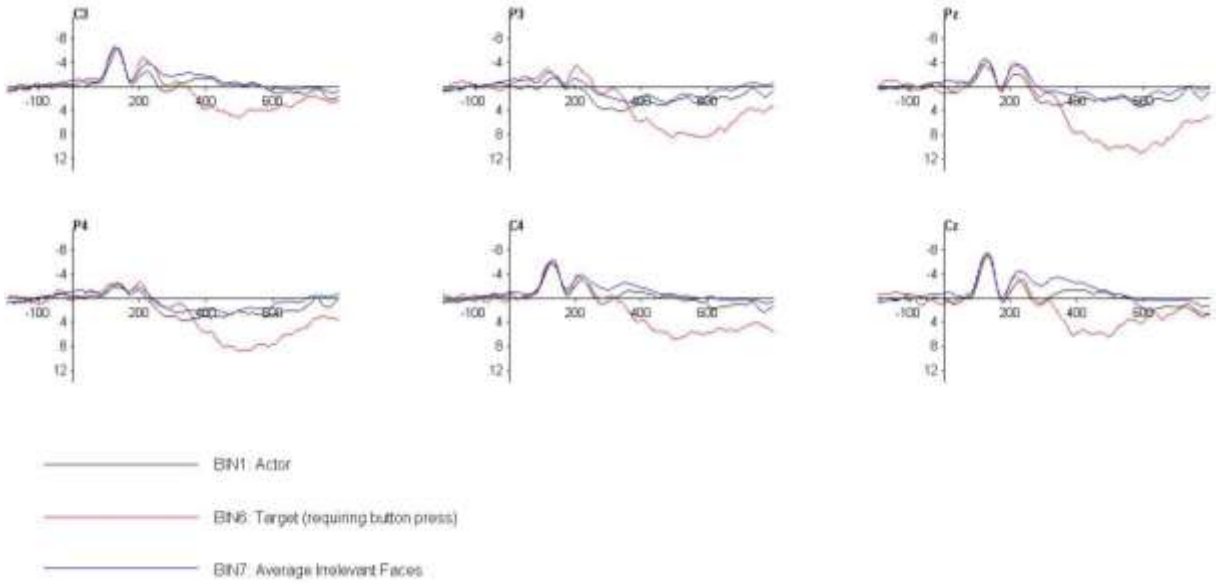

Figure 4: Famous experimental condition depicting grand average waveforms for the famous actor, the target requiring a button press, and the average famous irrelevant faces

Famous conditions ANOVAs

Contrary to the author's predictions for the famous conditions, a RM ANOVA comparing the P300 amplitude elicited for the famous actor across condition (famous control, famous experimental) and electrode site, showed no significant

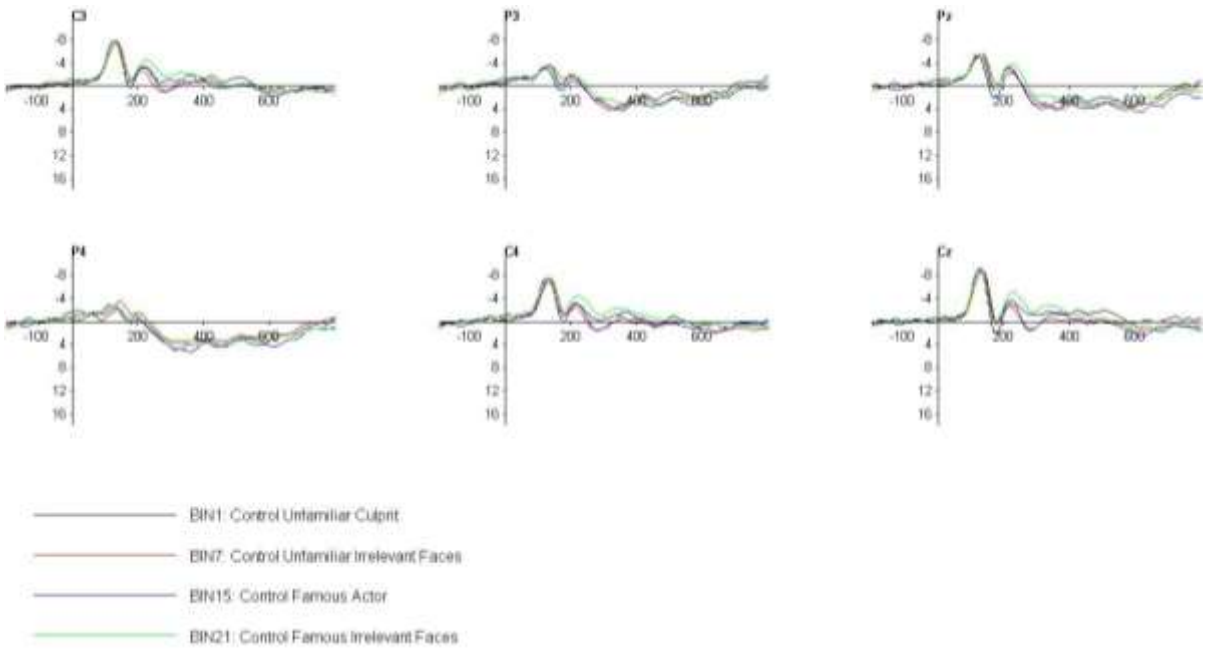

Figure 5: Grand average waveforms depicting the P300 of the unfamiliar culprit, unfamiliar irrelevant faces, famous actor, and famous irrelevant faces during the control conditions

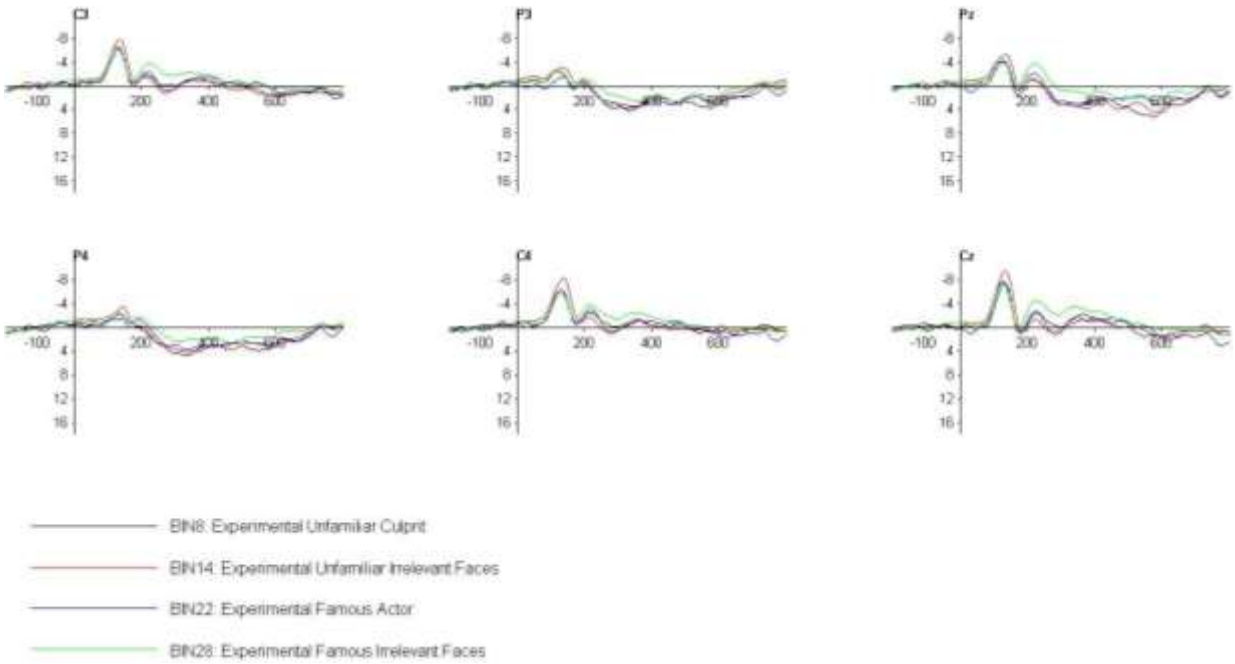

Figure 6: Grand average waveforms depicting the P300 of the unfamiliar culprit, unfamiliar irrelevant faces, famous actor, and famous irrelevant faces during the experimental conditions

difference in P300 amplitude elicited by the famous actor, between the famous control and the famous experimental condition, F (1, 11) = .18, $p > .05$.

Furthermore a RM ANOVA conducted on the famous experimental condition, comparing the P300 amplitude elicited by the famous actor and the famous

irrelevants across electrode sites, illustrated that participants displayed no significant difference in their P300 amplitude between the famous actor and famous irrelevant faces, F (1, 11) = 1.61, $p > .05$. This implies that participants did not recognise the famous actor both before and after watching the video in which he was present, no more than they recognised other famous faces.

Famous versus unfamiliar ANOVAs

Contradicting the authors predictions, a RM ANOVA comparing the unfamiliar culprit and famous actor in the control conditions, exposed no significant difference in the P300 amplitude, F (1, 11) = .60, $p > .05$.

A similar RM ANOVA comparing the culprit/actor across the unfamiliar experimental and famous experimental conditions, also found no significant difference in the P300 amplitude of the unfamiliar culprit and the famous actor in the experimental conditions, F (1, 11) = .17, $p > .05$.

Effects of the target on P300 amplitude

A larger RM ANOVA comparing across factors of condition (unfamiliar control, unfamiliar experimental, famous control, famous experimental), stimuli type (culprit/actor, target, irrelevant faces), and electrode site (C3, P3, Pz, P4, C4, Cz), found a significant effect of stimuli type, F (2, 22) = 45.91, $p < .000$. Additional post-hoc analysis revealed that the P300 amplitude for the target, which required participants to respond whenever they saw a repetition, was significantly larger than the culprit/actor ($p < .000$), and the irrelevant faces ($p < .000$) across conditions. This indicated that participants were paying attention to the stimuli.

Behavioural responses to line-up identification task

Figure 7 shows a graph depicting participant's responses to the line-up task. Specifically, the graph shows that participants were able to correctly identify the famous actor from the film clip 100% of time. In comparison, participants were able to correctly identify the unfamiliar culprit from the mock-crime video 50% of the time. Whilst trying to identify the unfamiliar culprit, 25% of participants reported that they did not know who they saw in the mock-crime video, and another 25% either identified the wrong individual, or did not think the culprit was in the line-up.

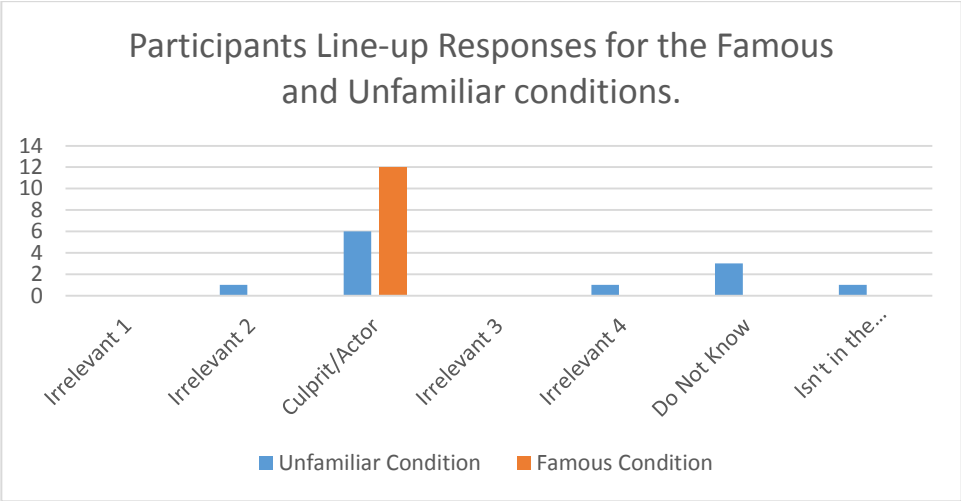

Figure 7: A graph displaying the identifications of the culprit or actor, made by participants after seeing the mock crime video or the film-clip

Discussion

Interpretation of unfamiliar face results

The main aim of this research was to provide evidence that poor recognition of unfamiliar faces was a factor in why eyewitness testimony can be inaccurate. It was predicted that an unfamiliar culprit would not be recognised more than other irrelevant unfamiliar faces after seeing the culprit in a mock-crime video. On inspection of the P300, results indicated participants were no better at recognising the culprit after seeing a mock-crime than beforehand, and that they recognised the culprit no better than anonymous irrelevant faces ($p > .05$). Inspection of the line-up identification rates showed that participants could only correctly identify the culprit at chance level.

This is consistent with the previous finding by the author that has shown how an unfamiliar culprit was not recognisable after seeing a mock-crime (Evans and Woolcock, unpublished). This is also consistent with the literature surrounding unfamiliar face recognition and matching ability (Hancock, Bruce, and Burton, 2000; Megreya and Burton, 2006), suggesting that the failure of participants to recognise the culprit neurologically, and also to identify the culprit in a line-up is likely due to the effect of poor-recognition of unfamiliar faces. A possible reason for this effect may be due to the low level of meaning an unfamiliar face would elicit in a participant. The P300 has been shown to activate to meaningful stimuli (Berlad et al, 1995; Gray et al, 2004; Johnson, 1993), therefore a lack of P300 elicited by unfamiliar faces may indicate that poor recognition for unfamiliar faces could be due to a lack of meaning; and thus improper encoding of memory for the unfamiliar face.

Interpretation of famous face comparisons

Another prediction of this study was that a famous actor will elicit a larger P300 compared to other famous faces after seeing the film-clip in which he is present, but not after watching a control video. In addition, it was also predicted that the famous actor would elicit a larger P300 than unfamiliar faces after watching both the control videos and the videos in which the culprit/actor were present. Contrary to predictions, the results indicated no significant difference between the famous actor and the famous irrelevant faces after watching a control video ($p > .05$). In addition to this, contradicting the author's predictions, results also showed no significant difference between the famous actor and the unfamiliar culprit both after watching control videos and after watching the mock-crime or film-clip ($p > .05$).

Interestingly, although there is no significant P300 wave which indicates recognition of the actor, the participants were consistently able to identify him in the line-up. Consistent with previous research on familiar/famous face recognition (Burton, Wilson, Cowan and Bruce, 1999; Ellis et al, 1979; Hancock et al, 2000), this indicates that participants did in fact recognise the famous actor.

This poses an interesting finding. There are several possible explanations for this result, possibly due to the methodology of the study. In the introduction, it was shown that Meijer et al (2009) was able to show a P300 due to mere recognition alone, by directing participant's attention elsewhere; an effect also shown by Gray et al (2004). However, this was in reaction to personally-familiar or self-relevant stimuli. Meijer, Smulders, Merckelbach and Wolf (2007) showed that when faces are known but personally less familiar, and no instructions are given to act on the specific face, it was harder to detect whether the participant recognised them based off their P300

response. This could be the case in the present study. Specifically, because famous faces are not very self-relevant or meaningful to the participant, and no specific instructions were given to act on upon the famous face during the ERP recordings, it is possible that this combination resulted in an ERP which was no different from unfamiliar faces.

Another possible explanation is that the results regarding the famous actor may well be due to the inclusion and presentation of the control condition. The control condition was presented first in both the unfamiliar and the famous condition. This gave participants no choice but to see the famous actor and the famous irrelevants roughly 35 times each before seeing the film clip in which the actor was present. Such a methodology may have desensitized participants to any 'rare recognised' effects which elicit a typical P300 wave, and although the faces are very well known, presenting a famous face amongst a sea of other famous faces before the experimental video could result in a habituation effect, in which famous, recognised faces become generic and expected, and then lose their meaning to the participant.

It's possible that these explanations may also explain the lack of P300 for the unfamiliar culprit. However, due to the low identification rates in the line-up for the unfamiliar culprit, it is more likely that the lack of P300 is due to the original explanation; in which participants simply couldn't recognise the unfamiliar face. Nonetheless, although the reasons for the results in the unfamiliar conditions seem likely, there is still a possibility that the results could be due to methodological errors, so must be interpreted gingerly.

Directions for future research

It is recommended by the author that future research utilizes this study and tries to improve on its methodology. Specifically, it is recommended that future research uses a between-subjects experimental design in which one group watch the control videos; and one group watch the mock-crime and film-clip videos. This will remove any confound provided by the possible habituation effect described earlier, although double the amount of participants will need to be recruited.

It is also recommended by the author that future research should instruct participants to act upon the measured stimuli in some way. As outlined (in the introduction) by Lefebvre et al (2007), participants could be instructed to respond to stimuli via a button press based on the category stimuli fall into. Examples of such categories may include: a button press for when the current picture on the screen depicts a culprit or actor they thought they saw in the video, or when the current picture on screen depicts an irrelevant unfamiliar or famous face that they thought was not in the video(s). It may also be useful for participants to respond to a target, such as when a picture repeats twice in a row on an infrequent basis; this can be used as a baseline, depicting a typical P300 ERP with which to compare other ERPs.

Less specifically, but nonetheless intriguing; it may be of use for future research to use stimuli at different angles or contexts. For example, research has indicated that unfamiliar face recognition is poor from different angles, lightings and presentations (Burton et al, 1999; Hancock et al, 2000). It would be rather interesting to find P300 and line-up results that vary as directional, lighting, and even episodic contexts change in the way that unfamiliar faces are presented.

Conclusions

Although some results contradict certain hypothesis made by the author, this study

provides some interesting implications. The present paper showed how participants could perform well on a line-up task identifying famous faces; but under the current methodology the P300 could not support this famous face recognition, possibly due to methodological issues. Importantly, the present paper found that poor performance on the eyewitness identification line-up task for an unfamiliar culprit was supported by a lack of P300 amplitude. Therefore, evidence is provided in this paper for the argument that inaccuracies found in eyewitness testimonies may well be due to a lack of recognition of unfamiliar faces; and this may be due to the unfamiliar face evoking little or no meaning to the eyewitness. Future research should make use of the guidance provided in this discussion section and try to improve on the methodology used in this study. This may possibly provide support for P300 recognition of familiar faces, as well as providing further evidence for the argument set out in this paper and clarification of any debatable results.

Acknowledgements

Firstly I would like to give my deepest thanks to my supervisor Dr. Giorgio Ganis. The advice and tutoring you have given to me over the past two years has been instrumental in my development as a scientist. I would like to thank David Bridges for your help, advice, and mostly your patience whilst teaching me how to set up the EEG lab equipment. To my family, Guy, Clare and Rhys; you are all amazing, thank you for every ounce of your encouragement, support and guidance. Without you I would not have come so far nor achieved so much. I'd like to give thanks to my housemates and friends, for the laughs we have shared helped keep me relatively sane throughout the development of this project. For all those who participated in this research, thank you; for without you there would not be a project for me to write acknowledgements for. Last but not least, I would like to thank Alex. Your support has been invaluable during the completion of this project. You have kept me healthy and happy. For that and much more, thank you.

References

Berlad, I., & Pratt, H. (1995). P300 in response to the subject's own name. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, 96 (5), 472-474.

Blackwood, D. H. R., Muir, W. J. (1990) Cognitive brain potentials and their application. The British Journal of Psychiatry; 157: 96–101

Brigham, J. C., & WolfsKeil, M. P. (1983). Opinions of attorneys and law enforcement personnel on the accuracy of eyewitness identifications. *Law and Human Behavior*, 7 (4), 337.

Buckhout, R. (1980). Nearly 2,000 witnesses can be wrong. *Bulletin of the Psychonomic Society*, 16 (4), 307-310.

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

Devine, D. J., Clayton, L. D., Dunford, B. B., Seying, R., & Pryce, J. (2001). Jury decision making: 45 years of empirical research on deliberating groups. *Psychology, Public Policy, and Law*, 7 (3), 622.

Ellis, H. D., Shepherd, J. W., & Davies, G. M. (1979). Identification of familiar and unfamiliar faces from internal and external features: some implications for theories of face recognition. *Perception*, (8), 431-9.

Gray, H. M., Ambady, N., Lowenthal, W. T., & Deldin, P. (2004). P300 as an index of attention to self-relevant stimuli. *Journal of Experimental Social Psychology*, 40 (2), 216-224.

Hancock, P. J., Bruce, V., & Burton, A. M. (2000). Recognition of unfamiliar faces. *Trends in cognitive sciences*, 4 (9), 330-337.

Johnson, R. (1993). On the neural generators of the P300 component of the event-related potential. *Psychophysiology*, 30 (1), 90-97.

Kebbell, M. R., & Milne, R. (1998). Police officers' perceptions of eyewitness performance in forensic investigations. *The Journal of Social Psychology*, 138 (3), 323-330.

Lefebvre, C. D., Marchand, Y., Smith, S. M., & Connolly, J. F. (2007). Determining eyewitness identification accuracy using event-related brain potentials (ERPs). *Psychophysiology*, 44 (6), 894-904.

Luck, Steven J. (2005). An Introduction to the Event-Related Potential Technique. The MIT Press.

Megreya, A. M., & Burton, A. M. (2006). Unfamiliar faces are not faces: Evidence from a matching task. *Memory & Cognition*, 34 (4), 865-876.

Meijer, E. H., Smulders, F. T., Merckelbach, H. L., & Wolf, A. G. (2007). The P300 is sensitive to concealed face recognition. *International Journal of Psychophysiology*, 66 (3), 231-237.

Meijer, E. H., Smulders, F. T., & Wolf, A. (2009). The contribution of mere recognition to the P300 effect in a concealed information test. *Applied psychophysiology and biofeedback*, 34 (3), 221-226.

Rosenfeld, J. P., Cantwell, B., Nasman, V. T., Wojdac, V., Ivanov, S., & Mazzeri, L. (1988). A modified, event-related potential-based guilty knowledge test. *International Journal of Neuroscience*, 42 (1-2), 157-161.

Sheck, B., Neufeld, P., & Dwyer, J. (2000). *Actual innocence*. New York: Random House.

The Justice Project, & United States of America. (2007). Eyewitness Identification: A Policy Review.

Wells, G. L. (1993). What do we know about eyewitness identification? *American Psychologist*, 48 (5), 553.

Wells, G. L., & Loftus, E. F. (2003). Eyewitness memory for people and events. In E. A. Goldstein (Ed.), Comprehensive handbook of psychology, Volume 11, *Forensic psychology*. New York: John Wiley and Sons.