THE USE OF EVENT-RELATED BRAIN POTENTIALS (ERPs) TO ASSESS EYEWITNESS IDENTIFICATION ACCURACY

by

Celeste D. Lefebvre

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

at

Dalhousie University
Halifax, Nova Scotia
October 2006

Your file Votre référence
ISBN: 978-0-494-27181-0
Our file Notre référence
ISBN: 978-0-494-27181-0

DALHOUSIE UNIVERSITY

To comply with the Canadian Privacy Act the National Library of Canada has requested that the following pages be removed from this copy of the thesis:

Preliminary Pages
- Examiners Signature Page (pii)
- Dalhousie Library Copyright Agreement (piii)

Appendices
- Copyright Releases (if applicable)

Dedication

L.C.B.

To the world you may be just one person, but to one person you may be the world.
~ Brandi Snyder

Thanks for all your support.

L.Y.A.

p.s. I am amazed by your attention to detail.

Table of Contents

List of Tables

List of Figures

Abstract

Mistaken eyewitness reports have consistently been found to be the major contributing factor leading to wrongful convictions. However, eyewitness testimony is regarded as important and strong evidence in the judicial system. Therefore, efforts are needed to increase the reliability of eyewitness reports. Using current lineup procedures, it is very difficult, if not impossible, to objectively determine if an eyewitness' identification is accurate or not. This thesis investigated the use of event-related brain potentials (ERPs) as a potential tool to provide an objective measure of eyewitness accuracy during a lineup identification task (ERP-lineup). In addition, this thesis also aimed to ascertain the neurophysiological impact of several variables that have been demonstrated to influence eyewitness accuracy. Experiment 1 assessed whether the photograph of the criminal elicited a specific ERP response compared to the other lineup members (standard condition) as well as the impact of deliberate attempts to conceal the identity of the criminal (deception condition). Experiment 2a assessed the impact of the time delay between when participants viewed the crime until when they completed the ERP-lineup (no-delay, 1-hour delay and 1-week delay conditions). Lastly, Experiment 2b examined the impact of whether the photograph of the criminal was in the ERP-lineup (criminal-present condition) or not (criminal-absent condition). Taken together, the results from Experiments 1, 2a and 2b, demonstrated that a centro-parietal late positive complex (LPC, elicited between 400 and 600 ms post-photograph onset) provided a neurophysiological index of explicit recognition of the criminal. This effect remained strong, irrespective of the time delay between seeing the crime video and the ERP-lineup task or whether participants attempted to deny recognition of the criminal. In addition, the LPC was attenuated or was not elicited when the criminal was absent from the lineup. Although more research is needed before an ERP-lineup task should be applied to real-world cases, the results are promising and warrant continued research.

List of Abbreviations and Symbols Used

ANS = Autonomic nervous system

C = Criminal

CA = Criminal-absent

CI = Correct identification (selection of the criminal from a criminal-present lineup)

CIT = Concealed information task

CP = Criminal-present

CR = Correct rejection (selection that the criminal is absent in a criminal-absent lineup)

DNA = Deoxyribonucleic acid

EEG = Electroencephalogram

EOG = Electrocculogram

ERN = Error-related negativity

ERP = Event-related brain potential

F1 = Filler 1

F2 = Filler 2

F3 = Filler 3

F4 = Filler 4

F5 = Filler 5

FI = False identification (selection of the suspect in a criminal-absent lineup)

FR = False rejection (selection that the criminal is absent in a criminal-present lineup)

GKT = Guilty knowledge test

ID = Identification decision

IS = innocent suspect (the person suspected of committing the crime is actually innocent)

LPC = Late positive complex

M = Mean

MI = Misidentification (selection of a filler in either a criminal present or absent lineup)

ms = milliseconds

RM ANOVA = Repeated measures analysis of variance

SD = Standard deviation

V = Victim

kΩ = kilolhms

μV = microvolts

Acknowledgements

My Supervisor: John Connolly

The whole art of teaching is only the art of awakening the natural curiosity of young minds. ~ Anatole France
Thanks for providing me with the opportunity and encouragement to follow my curiosity.

My Committee Members:

To get something done, a committee should consist of no more than three, two of whom are absent ~ Robert Copeland
Nonsense! Thanks for your input.

Bradley Frankland
In ancient times they had no statistics so they had to fall back on lies. ~ Stephen Leacock
With your help, I have been able to use statistics to aid in detecting lies.

Steven Smith
The secret of a good memory is attention, and attention to a subject depends upon our interest in it. We rarely forget that which has made a deep impression on our minds. ~ Tryon Edwards
Irrespective of the time delay, I will always remember Simon and Peter. In addition to your advice, thanks a lot for the stimuli.

My External Examiner: J. Peter Rosenfeld

Without lies humanity would perish of despair and boredom. ~ Anatole France
Your work on lie-detection with ERPs was an inspiration for me. It was an honour to have you as my external examiner.

Yannick Marchand:

Some cause happiness wherever they go; others, whenever they go. ~ Oscar Wilde
You are clearly the former!

Lab Members and Volunteers: Paula Miles, Troy Harker, Jeannine Gravel, Danielle Bourque, Gerry McCormick, Vanessa Versteg, Sarah Applebaum and Rachel Clay

The only people with whom you should try to get even are those who have helped you. ~ John E. Southard
Thanks for all your assistance, advice and hard work.

Friends: You know who you are!

Quit now, you'll never make it. If you disregard this advice, you'll be halfway there. ~ David Zucker
Friends really do make all the difference! Thanks for getting me through the times when I doubted that I could do it. And, of course, thanks for all the fun!

My Life-Long Friends: Agata Kowalski, Vicki Greene and Kelly Maslen

When I think back to the "good times", the memories are of you. ~ Anonymous

Family:

Don't take life too seriously, you'll never get out of it alive. ~ Elbert Hubbard
You have always been there to make sure that I don't take myself too seriously.

… I hope that I have accurately identified all.

Chapter 1: General Introduction

Wrongful convictions and eyewitness testimony

Within the last decade, the use of deoxyribonucleic acid (DNA) analysis has been applied to past criminal cases for which DNA-rich evidence was preserved. The results of these analyses have revealed that a number of innocent individuals were found guilty and punished for crimes they did not commit. By 1998, DNA evidence revealed that over forty cases of wrongful convictions had occurred in the United States (Connors, Lundregan, Miller, & McEwen, 1996; Wells, et al., 1998) and on average, these individuals spent 8.5 years in jail and five were sentenced the death penalty (Wells, et al., 2000). Most strikingly, in 90% of these cases, faulty eyewitness testimony was the major factor leading to the unjust guilty verdict. Since these reports, an additional 62 individuals have been exonerated based on DNA analyses; eight of whom were sentenced the death penalty (Scheck, Neufeld, & Dwyer, 2000). Eyewitness errors accounted for 84% (52/62) of these cases. Furthermore, there have been 36 documented cases of wrongful convictions as a result of mistaken eyewitness reports in Great Britain (Devlin, 1976) and there have been numerous reports from Canada on a case-by-case basis (e.g., Regina v. Sophonow 1986; Regina v. McGuinnis, Ballantyne & Ballantyne1997).

Overall, mistaken eyewitness reports have consistently been found to be the major contributing factor leading to wrongful convictions (Brandon & Davies, 1973; Connors, Lundregan, Miller, & McEwen, 1996; Huff, Rattner, & Sagarin, 1986; Loftus, 1979; Wells, et al., 1998; Woocher, 1977), and in fact are more common than all other reasons combined (Wells & Olsen, 2003). Speculation regarding eyewitness inaccuracies is not new and reports date back to the early 1900's (Borchard, 1932; Brandon & Davies, 1973;

Connors, Lundregan, Miller, & McEwen, 1996; Frank & Frank, 1957; Huff, Rattner, & Sagarin, 1986; Münsterberg, 1908).

It is estimated that in the United States there are over 77,000 suspects placed in lineups every year (Wells, et al., 1998) and that approximately 200 people are identified from live or photo lineups every day (Goldstein, Chance, & Schneller, 1989). These statistics highlight both the frequency of lineup identifications and the serious impact of eyewitness procedures. The importance of this is even more apparent considering that in some countries the testimony of only one eyewitness is sufficient to lead to a conviction (Wright & Stroud, 2002). Although the actual rates of falsely convicting an innocent individual in court cases is unknown, estimates from the United States and Great Britain suggest, that for serious crimes, one to 10 out of every thousand people are wrongfully convicted (Rosen, 1992). Yet, data from the Federal Bureau of Investigation's own laboratories suggest that the rate may be much higher. Of 8,000 suspects arrested for sexual assault who had DNA samples tested by the US Federal Bureau of Investigation laboratory, over 2,000 (or about 25%) were excluded as the perpetrator. Typically these arrests were made based (at least in part) on faulty eyewitness identification (Scheck, Neufeld, & Dwyer, 2000).

Why is eyewitness testimony important?

Studies investigating the factors that jurors, judges, police, prosecutors and other law enforcement personnel consider most pertinent for determining guilt from innocence have provided support that eyewitness testimony is considered to be very strong evidence (Brigham & WolfsKeil, 1983; Cutler, Penrod, & Stuve, 1988; Wells, 1984b). Generally, most types of evidence presented in court are circumstantial or fail to offer a direct link between the suspect and the execution of the criminal act. For example, physical trace evidence such as fingerprints, blood samples or DNA can only place the suspect at the scene of the crime (with the exception of semen DNA that can directly implicate the individual in rape cases). When an eyewitness selects an individual from a lineup, they are effectively providing a direct and powerful link that the suspect did in fact commit the crime (Wells & Loftus, 2002; Wells, et al., 2000). Additionally, most crimes (with the exception of semen samples from sexually-based offences) do not involve DNA-rich biological traces and therefore the judicial system remains heavily reliant on eyewitness identification (Wells & Olsen, 2003). Thus, eyewitness testimony is an important form of evidence.

Disregarding or dismissing eyewitness reports could be potentially devastating, particularly for serious and dangerous offenses. For obvious reasons, when a serious crime occurs, there is pressure to solve the case because of the risk the criminal poses to the community. In the judicial system, there is debate as to which is the greater failure – to convict a potentially innocent person or to allow a dangerous offender go free. For example, a dilemma exists on what to do when an eyewitness identifies a suspect but there is a lack of strong physical evidence. Overall, it is important to decrease any type of error within the system. Therefore, despite biases and the risk of mistaken identifications, eyewitness testimony remains important and can offer unique and crucial evidence. It seems obvious that rather than completely dismissing eyewitness reports, efforts should go toward increasing their reliability.

There is an apparent discrepancy within the legal system regarding its treatment of physical trace evidence compared to eyewitness or memory-based evidence (Wells &

Loftus, 2002). The collection of physical trace evidence involves relatively well-formalized scientific protocols and is predominantly conducted by forensic science specialists. By contrast, law enforcement personnel typically collect eyewitness evidence with little or no training in human memory. Wells and Loftus (2003) believe that the rapid advances in the preservation of physical trace evidence have been a result of the adoption of the scientific model. One the other hand, these authors believe that the lack of application of laboratory-based eyewitness and memory research findings is at least partially responsible for some eyewitness errors. These authors also argue that well-formalized protocols based on scientific study should be developed and applied to prevent contamination or loss of memories in an analogous manner as for physical trace evidence (Wells & Loftus, 2003).

Eyewitness identification procedures

Within the last few decades, there has been extensive investigation into eyewitness testimony and variables that can alter accuracy rates. By 1995, over 2000 scientific investigations were documented in the research literature on eyewitness identification (Cutler & Penrod, 1995). Although there is now extensive evidence that eyewitness testimony is prone to errors and biases that can lead to mistaken identifications, a large body of eyewitness research is targeted at improving eyewitness accuracy.

A predominant area of eyewitness research is to determine ways of decreasing avoidable errors by manipulating factors that are under the control of the justice system [referred to as system variables (Wells, et al., 2000)]. Two ways of avoiding such errors has been to develop strategies/procedures that 1) discourage the use of a relative judgment strategy (i.e., to decrease the likelihood that an eyewitness will select a lineup member based solely on the fact that they *look the most like the criminal* or 2) to create non-biased lineups to prevent the suspect for standing out in a lineup (Lindsay, Lea, Nosworthy, & Fulford, 1991; Wells, 1984c; Wells, et al., 2000; Wells & Olsen, 2003).

In 1998, a panel in the United States was formed at the request of the Attorney General of the U.S. Department of Justice, National Institute of Justice, to establish national best practice guidelines for the administration of live and photograph lineups. In 1999, a guidebook was published outlining recommendations based on the decades of eyewitness research investigating system variables (Reno, Fisher, Robinson, Brennan, & Travis, 1999). This guide has subsequently been updated in 2003 (Ashcroft, Daniels, & Hart, 2003). The goal of the guidelines was to decrease the two classes of system biases (i.e., decreasing reliance on relative judgment strategies and ensuring the use of non-biased lineup procedures). For simplicity and an attempt to standardize lineup procedures, the guidelines were reduced to four major Rules.

The four major Rules include 1) double-blind administration of the lineup (i.e., the identity of the suspect is unknown to the administrator of the lineup, and the person who constructs the lineup is not the administrator). 2) Lineup instructions should clearly warn that *the criminal might or might not be present in the lineup.* 3) The additional members of the lineup (referred to as fillers) should be selected based on overlapping features with the criminal or from eyewitnesses' verbal description of the criminal in terms of approximate height, weight, hair features, age and gender. 4) Confidence ratings regarding eyewitnesses' selection should be documented immediately (Piggott & Brigham, 1985; Wells, 1985).

In addition, although not an explicit rule, the guide states that only one suspect is to be used per identification procedure, and if there are multiple suspects, a separate lineup should be created for each (Ashcroft, Daniels, & Hart, 2003 , p. 33). To clarify, the term *suspect* should not be confused with *criminal* or with *fillers*. A suspect might or might not be the criminal and the fillers are *known innocents* and therefore are not suspects (Wells & Loftus, 2003). If these Rules and guidelines are followed, then misinterpretations should, ideally, only arise from the eyewitness' recognition memory, rather than from biases attributable to system variables.

The four Rules are based on a number of key factors that have been found to impact eyewitness accuracy. The double-blind recommendation is based on the known cognitive phenomenon that people have a natural propensity to try to confirm a hypothesis and hence may consciously or unconsciously bias the evidence (Fischhoff & Beyth-Marom, 1983; Nickerson, 1998). Knowledge of the suspect in the lineup by administrators has been shown to have an impact on eyewitness identifications. For example, smiling and nonverbal reinforcement by the lineup administrator during the presentation of the suspect can lead to a selection bias (Fanselow & Buckout, 1976). Field studies have also indicated that lineup administrators say things that focus the eyewitness on the suspect (Wells & Seelau, 1995). The major rationale for this rule is that if a lineup administrator is blind to the suspect then they should not be susceptible to biasing the eyewitness.

Rules 2 and 3 stem from the relative judgment phenomenon - that people select the lineup member that looks most like the culprit (Wells, 1984c). This problem amplifies when the suspect in the lineup is innocent because he/she was selected based on overlapping features and a resemblance with the actual criminal. For Rule 2, the empirical data demonstrate that an eyewitness is less likely to identify an innocent suspect when they are warned that the criminal might not be present in the lineup (Malpass & Devine, 1981; Parker & Carranza, 1989; Parker & Ryan, 1993). The use of the *might or might not be present in the lineup* clause is believed to reduce mistaken identifications by decreasing reliance on the relative judgment process but also by legitimizing the behaviour of not selecting anyone.

For Rule 3, all the members of the lineup should resemble the suspect in physical features. If a suspect is the only person that fits the eyewitness' general verbal/physical description, this acts to disproportionately increase the bias toward the suspect. However, if all the lineup members share at least some basic physical features with the suspect, then relative judgment errors should be equal across all the lineup members. Therefore the use of Rules 2 and 3 serve to increase the likelihood of an absolute judgment process and, thus, the eyewitness' identification is more likely to be based predominantly on their memory of the criminal event.

In laboratory studies, mock jurors tend to overestimate eyewitness' accuracy rates (Lindsay, Lim, Marando, & Cully, 1986; Lindsay, Wells, & O'Connor, 1989a; Wells & Leippe, 1981). Currently, an eyewitness' level of confidence is the most powerful determinant of whether or not observers believe that the eyewitness' decision is accurate or not (Cutler, Penrod, & Dexter, 1990; Leippe & Romanczyk, 1989; Lindsay, Wells, & O'Connor, 1989b; Wells & Leippe, 1981; Wells & Olsen, 2003). Rule 4 is based on the empirical support that eyewitnesses' confidence ratings can become inflated from the time of the identification until the testifying date. Of great concern, there is also the evidence that witnesses are able to extract and incorporate new information after witnessing an event and then testify as if the new information was part of the original event (Loftus, 1979; Wright, 1993). In addition, multiple system variables have been shown to impact eyewitnesses' confidence levels (usually resulting in confidence inflation), irrespective of eyewitness' accuracy. Some system factors shown to alter confidence include: asking misleading questions (e.g. Loftus & Loftus, 1980), increasing the number of times an eyewitness is questioned (e.g. Hastie, Landsman, & Loftus, 1978; Turtle & Yuille, 1994) practicing cross-examination questions (Wells & Leippe, 1981) and knowing the identification decision of a co-witness (Luus & Wells, 1994). Therefore, failing to take confidence ratings at the time of identification can be problematic.

Two additional procedures are also highly recommended but not listed as formal rules. Firstly, to use sequential lineups opposed to the more routine practice of simultaneous lineups. Simultaneous lineups involve presenting multiple photographs to the eyewitness at the same time. This allows the eyewitness to compare across the lineup members before making their decision. In contrast, for sequential lineups, only one photograph is presented at a time. This approach requires the eyewitness to make a decision about each photograph prior to the presentation of the next lineup member. This recommendation is based on evidence that the use of a sequential procedure significantly improves the rate of correct rejections and decreases false identifications in criminal-absent (CA) lineups. In addition, this procedure has a minimal or negligible impact on the correct identification rate for criminal-present (CP) lineups (Lindsay, Lea, & Fulford, 1991; Lindsay, Lea, Nosworthy, & Fulford, 1991; Steblay, Dysart, Fulero, & Lindsay, 2001). Secondly, videotaping the lineup procedure is recommended to ensure that the guidelines are followed appropriately.

Sadly, despite the establishment of the research-based guide book of recommendations in 1999, it was estimated that out of 13,000 police departments sampled in the United States, only a few altered their policies based on the guide (Wells, et al., 2000) and that there was a great deal of disparity across the states, jurisdictions and even between individual police officers within the same service (Turtle, Lindsay, & Wells, 2003). However, one study reported a high rate of compliance with similar research-based recommendations in the United Kingdom (Kebbell, 2000).

Canadian Guidelines have also been established. In 1983, Neil Brooks (a law professor) was hired by The Law Reform Commission of Canada to develop guidelines with the intent of 1) increasing the reliability of identification procedures 2) to reduce the risk of mistaken identification, and 3) to protect the rights of suspects (Brooks, 1983). Several prominent Canadian and American eyewitness researchers acted as consultants. Thirty-eight recommendations were devised. In 2001, Supreme Court Justice Cory conducted a public inquiry regarding the wrongful conviction of Thomas Sophonow, 1986 (*Regina vs. Sophonow)*. The resulting report specified 48 recommendations, 16 of which pertained to eyewitness identification procedures (Cory, 2001). In both of the Canadian reports the four Rules discussed above were included [refer to Yarmey (2003), for an overview of these guidelines].

Eyewitness identification accuracy

Retrospective field studies

Generally, it has been found that mistaken identifications and lineup rejections are not well-documented (Tollerstrup, Turtle, & Yuille, 1994). However, there have been a few retrospective studies that have attempted to investigate identification decisions based on real criminal cases. A study conducted in the United Kingdom found that 20% of eyewitnesses selected a filler from lineups (Wright & McDaid, 1996). The most comprehensive archival study based on real crimes conducted to date examined 279 crimes, involving 258 showups [i.e., the eyewitness is provided with an opportunity to view the suspect (live or in a photo) without other lineup members present], 289 photographic lineups and 58 live lineups. The results showed that the suspect was identified in 76% of showups and in 48% of photographic lineups. For live lineups, the procedures were more carefully documented, which allowed for calculating the frequency of individuals that selected one of the fillers and the rates of lineup rejections (i.e., incorrectly stating that the criminal was not present in the lineup) rates. For live lineups, 50% selected the suspect, 24% selected a filler and 26% were not able to make a choice. The identification rates were significantly higher for showups compared to live and photographic lineups. This pattern was expected because an identification made during a showup is always of the suspect. In addition, 93% of showups occurred within one day of the crime, whereas, only 55% of photographic lineups occurred within the first week following the crime (the time delay information was inaccessible for live lineups). Therefore, it is also possible that the shorter time delay may have also contributed to the higher identification of suspects during showups (Behrman & Davey, 2001).

Controlled laboratory studies

There have been three recent meta-analyses that have investigated eyewitness accuracy rates. Firstly, in a meta-analysis of 48 experiments, the overall rate of correct identifications in CP lineups was 51%. Twenty-seven percent made misidentifications (i.e., incorrectly selected one of the fillers) and 21% made false rejections (i.e., incorrectly stated that the criminal was absent from the lineup). For CA-lineups, the correct rejection rate (i.e., correctly stated that the criminal was absent from the lineup) was 43% and the misidentification rate was 57% (Haber & Haber, 2001).

In a second meta-analysis, involving 23 controlled studies with 4,145 participants, correct identifications were 35% for sequentially presented CP lineups, with 46% false rejections and 19% misidentifications. Rates for simultaneous CP lineups were 50% correct identifications, 26% false rejections and 24% misidentifications. For sequentially presented CA lineups, 72% made correct rejections (i.e., correctly stated that the criminal was absent from the lineup) and 28% made misidentifications for lineups. However for simultaneous CA lineups, 49% made correct rejections and 51% made misidentifications. Collapsing across lineup type (CP and CA), the overall correct decision rate (referring to either a correct identification in a CP lineup or a correct rejection in a CA lineup) was 56% for sequential lineups and 48% for simultaneous lineups (Steblay, Dysart, Fulero, & Lindsay, 2001).

Lastly, the third and most recent meta-analysis involved eight studies and 3,013 participants. The lineup procedures were comparable across the eight studies selected (e.g., all experiments used one suspect lineups with five fillers and warning instructions that the criminal might or might not be present in the lineup). For CP lineups, correct identifications were 45%, misidentifications were 24% and false rejections were 34%. For CA lineups, 57% made correct rejections, and 43% made misidentifications. Overall, correct decisions were 51% collapsed across the CP and CA lineups (Steblay, Dysart, Fulero, & Lindsay, 2003).

Ecological validity of controlled laboratory studies

A great deal of controversy exists over how frequently false identification errors occur in real-world cases and how well one can generalize from laboratory studies to the field (Loftus, 1983; Wells, 1984a). Most people would like to believe that misidentification rates demonstrated in the laboratory (which range from 20-80%) are much higher than what occurs in the real world. This supposition is based on the belief that witnessing or being the victim of an actual crime involves an increased level of intensity and personal relevance (and possibly personal threat) as well as legal formalities and consequences that are obviously not present during laboratory studies. However, in contrast with real-world crime situations, witnessing conditions in laboratory experiments offer 1) ideal/adequate viewing conditions 2) shorter retention intervals (e.g., a few hours/days compared to a few months or even possibly years in real-world cases) 3) use of college age students (whose mean age, health, and memory abilities are better than eyewitnesses from the general population) and 4) decreased stress. All of these factors are associated with enhanced memory retrieval (Bothwell, Brigham, & Pigott, 1987; Deffenbacher, Bornstein, Penrod, & McGorty, 2004; Morgan, et al., 2004; Peters, 1988; Southwick, Morgan, Nicolaou, & Charney, 1997). In addition, laboratory studies and staged crimes generally provide an ecologically satisfactory portrayal of crimes such as robbery and theft. Therefore, one can argue that the accuracy rates found in laboratory experiments most likely *under* represent eyewitness identification error rates in real cases (Wells & Turtle, 1986).

An objective marker of accurate identification?

Currently, it is very difficult to objectively determine whether an eyewitness' identification of a suspect is accurate or not. One new avenue is to investigate eyewitness accuracy by using event-related brain potentials (ERPs) with the goal of providing a reliable neurophysiological index of an eyewitness' recognition memory. In other words, is there a certain brainwave pattern that is elicited when an eyewitness is presented with and recognizes the face of the perpetrator compared to the fillers in the lineup?

Event-related brain potentials (ERPs) are acquired from electroencephalographic (EEG) recordings and provide a non-invasive, temporally accurate on-line index of the brain's electrical activity in response to a stimulus or cognitive event (Chiappa, 1997; Knight, 1997; Regan, 1989). Changes in voltage distribution across the scalp over time are presumed to reflect the activation and orientation of neuronal ensembles within the brain. Through a process of signal averaging, the electrical activity related to the processing of multiple presentations of a stimulus (or a class of related stimuli) is retained (Brandeis & Lehmann, 1987). An ERP component refers to a peak or trough in the averaged waveform (or some other reliable type of EEG change) that is associated with the cognitive, sensory or motor processing of a stimulus. A component is generally defined by its polarity (positive, P, or negative, N), peak latency (post-stimulus onset) and topography (electrode sites where the potential magnitude is maximal). Since the ERP signal averaging technique has been established, this method has been used extensively to investigate cognitive functioning in both healthy and patient populations.

Currently there have been no ERP studies conducted using an eyewitness lineup identification paradigm, however there is now a large body of ERP literature investigating the cognitive processing associated with recognition memory as well as the use of ERPs in other forensic contexts (e.g., for the detection of deception or malingering).

ERPs and recognition memory judgments

A recognition memory judgment refers to the decision regarding whether a current event corresponds to a previously experienced event. There are two general theories about how decisions are made about whether something is remembered or not, either the judgment is comprised of 1) a unitary process that is determined by memory strength (Donaldson, 1996; Murdock, 1965) or 2) a dual-process, whereby events can be recognized as either *familiar* or as a *recollection*. Familiarity refers to the sense of having seen an item before and recollection refers to remembering an item along with retrieval of perceptual, contextual or source information (see Yonelinas, 2002). Although, there is still some controversy surrounding these two theories, currently most evidence supports the dual-process theory of recognition memory judgments.

The extensive literature on recognition memory performance has consistently demonstrated that ERPs elicited to old items (previously studied stimuli) are more positive than to new items (novel, not previously presented stimuli) within 300-1000 ms post-stimulus onset (see Friedman & Johnson, 2000; Mecklinger, 2000; Rugg, 1995). The positivity elicited to old items has been broken down into two main temporal windows that are believed to correspond to two different and topographical components, 1) the frontal (or early, 300-500 ms post-stimulus onset) and 2) the parietal (or late, 400-1000 ms post-stimulus onset) old/new effects (Rugg, Allan, & Birch, 2000; Schloerscheidt & Rugg, 2004). The frontal and parietal labels refer to the scalp locations where the potentials' positive voltage is maximal, rather than to frontal and parietal brain structures.

The frontal old/new effect is characterized by a fronto-central positivity elicited to old items within the 300-500 ms latency range compared to new items. On the other hand, the parietal old/new effect is characterized by a prolonged positive shift between 400-1000 ms which is maximal over the parietal scalp to old items in contrast new items (see Friedman & Johnson, 2000; Mecklinger, 2000; Rugg, 1995). In line with the postulated dual-processing of recognition memory, the frontal old/new effect is generally considered to be associated with familiarity (a sense of having previously encountered the item), meanwhile the parietal old/new effect is postulated to underlie recollection [explicit recognition and retrieval of item specific information (Hintzman, Caulton, & Levitin, 1998; Reder, et al., 2000; Yonelinas, 1994; Yonelinas, 1999; Yonelinas, 2002; Yonelinas, Dobbins, Szymanski, Dhaliwal, & King, 1996)].

According to leading models of recognition memory judgments, familiarity results from a matching process that compares the global similarity from test items to previously studied items (Clark & Gronlund, 1996; Gillund & Shiffrin, 1984). Although the frontal old/new effect has been shown to distinguish old versus new items when they are dissimilar from each other, it has failed to reliably differentiate old items from new, highly similar (lure) items (Curran, 2000; Curran, Tanaka, & Weiskopf, 2002; Nessler, Mecklinger, & Penney, 2001). In addition, other studies have found that participants are able to distinguish old from new items (based on corresponding old/new button press responses made following stimuli presentation during an ERP task) after intervening words or retention intervals of 15-45 minutes. However, the frontal old/new effect is not maintained across these longer time intervals (Rugg, 1990; Rugg & Nagy, 1989; Rugg, Rugg, & Coles, 1995). This suggests that the frontal old/new effect may be fleeting and have limited impact on longer-term memory processes. However, one study using picture stimuli was able to demonstrate the maintenance of both the frontal and parietal old/new effects following 34-minute, 39-minute and 1-day retention intervals. It may be that picture stimuli are more resistant to forgetting than words (Curran & Friedman, 2004). Overall, the association of the frontal old/new effect to a sense of familiarity, rather than recollection, suggests that it may not be able to distinguish the criminal from other lineup members, which are generally selected based on similarities and overlapping features with the criminal.

With regard to the parietal old/new effect, there is a general consensus within the literature that it is correlated with the recollection of old items (Paller & Kutas, 1992; Rugg, Cox, Doyle, & Wells, 1995; Rugg, et al., 1998; Smith & Guster, 1993) and indexes explicit recognition memory beyond familiarity (e.g. Duzel, Yonelinas, Mangun, Heinze, & Tulving, 1997; Rugg, et al., 1998; Ullsperger, Mecklinger, & Muller, 2000). Firstly, the parietal old/new effect has been demonstrated to be sensitive to recognition accuracy, as it is elicited only to correctly recognized stimuli but not to incorrectly classified items (Curran, 1999; Rugg, 1995; Rugg, et al., 1998).

Secondly, a remember/know paradigm has also been used to dissociate *remember* (explicit recollection) versus *know* (sense of familiarity) subtleties and has found that the parietal old/new effect is larger to recollection compared to familiarity-based responses (Duzel, Yonelinas, Mangun, Heinze, & Tulving, 1997; Rugg, Schloerscheidt, & Mark, 1998). Thirdly, the parietal old/new effect has been found to be larger (and the frontal old/new effect to be unaffected) when deeper levels of processing are imposed (Paller & Kutas, 1992; Rugg, Schloerscheidt, & Mark, 1998) and when tasks demand the retrieval of contextual information (Rugg, Schloerscheidt, & Mark, 1998; Schloerscheidt & Rugg, 2004; Wilding & Rugg, 1996), both of which are factors presumed to be more associated with recollection than with familiarity. Hence, the parietal old/new component has been proposed to be a potential neural correlate of explicit recognition memory.

ERPs and recognition memory for faces

The parietal old/new effect has also been investigated in face recognition paradigms and has consistently been demonstrated to be elicited to famous or family photographs compared to non-famous faces (e.g. Barrett & Rugg, 1989; Barrett, Rugg, & Perrett, 1988; Bentin & McCarthy, 1994; Guillem, Bicu, & Debruille, 2001; Hepworth, Rovet, & Taylor, 2001; Nessler, Mecklinger, & Penney, 2005; Paller, Bozic, Ranganath, Grabowecky, & Yamada, 1999; Schweinberger, Pfütze, & Sommer, 1995). In addition, the parietal old/new effect is more pronounced in explicit face memory tasks compared to implicit ones (Paller, Bozic, Ranganath, Grabowecky, & Yamada, 1999; Paller, Gonsalves, Grabowecky, Bozic, & Yamada, 2000) and it is elicited to correctly identified old faces and not to faces incorrectly categorized as old (Joyce & Kutas, 2005). The parietal old/new effect has also been found to be equally elicited to correctly identified repetition of both famous and non-famous faces, providing further support for a link between the parietal old/new effect and explicit recognition accuracy (Nessler, Mecklinger, & Penney, 2005).

As already discussed above, the parietal old/new effect is not face-specific. One study directly contrasted the waveforms for famous and non-famous faces and names (Schweinberger, Pickering, Burton, & Kaufmann, 2002). The parietal old/new effect was equally elicited to famous faces and the names of famous people and scalp distribution patterns generated by names and faces were virtually indistinguishable, suggesting similar brain sources.

Interestingly, one facial recognition study (Joyce & Kutas, 2005) compared newly learned faces to novel faces and found that the parietal old/new effect was larger following a longer retention interval (i.e., amplitude was higher at 1-day and 1-week intervals compared to no-delay, 30-minute and 1-hour delays). Generally, it is believed that recognition accuracy decreases following longer time delays due to forgetting, and therefore one would expect a reduction of positive amplitude as recognition accuracy decreased. However, in this study, only correct responses were included in the ERP averages. Therefore, it could be that the photographs remembered at 1-day and 1-week intervals had a stronger memory trace, which is indexed by the higher amplitude. Alternatively, the larger amplitude may also be related to additional processing or active reconstructing of the memory trace that may be necessary when the memory occurred further in the past. If either interpretation is correct, then it lends encouraging support for the use of the parietal old/new effect as a gauge of explicit recognition memory overtime. Thus, it suggests that the use of ERPs in an eyewitness context could be used in the field, where retention intervals are variable and often long.

Face recognition versus eyewitness identification tasks

A recent meta-analysis (involving non-ERP studies) was conducted to review and synthesize findings on face and identity recognition (Deffenbacher, Bornstein, Penrod, & McGorty, 2004). The authors conducted separate analyses for studies using an eyewitness paradigm (i.e., presentation of a crime scenario followed by a lineup task) compared to a face recognition paradigm [i.e., presentation of a set of faces (study phase) followed by a recognition task comparing faces from the study phase embedded among new faces]. Generally, the face recognition studies had a larger number of target faces (e.g., 24 or more). However the percentage of distracters was less (e.g., 50% new and 50% old faces) compared to eyewitness paradigms (e.g., five or more distracters for each target). When comparing these two types of paradigms, accuracy rates tended to be substantially lower for eyewitness contexts. It was inferred that the higher stress levels associated with a crime-related context accounted for the decreased accuracy in the eyewitness studies. Correct identifications were found to be 56% for facial recognition studies compared to only 39% for high stress eyewitness paradigms. However, for low stress eyewitness designs, correct identifications rates were more comparable (53%) to rates for face recognition paradigms (58%). Overall, stress or interrogative recall commonly associated with an eyewitness paradigm was found to have a negative impact on facial recognition (Deffenbacher, Bornstein, Penrod, & McGorty, 2004). Therefore, there is reason to believe that there may also be ERP differences associated with an eyewitness paradigm compared to more standard facial recognition tasks. This highlights the need for ERP studies using eyewitness protocols.

The use of ERPs in forensic contexts

In addition to ERP research on recognition memory, there has also been interest in the use of ERPs in a forensic context to assist or facilitate the detection of deception. The application of ERPs to deception detection involves investigating an individual's neural activity while they are presented with crime-relevant and crime-irrelevant stimuli. The goal is to determine if the crime-related information can be reliably differentiated from the crime-irrelevant information in people with knowledge of the crime. The premise is that a perpetrator (or accomplices and eyewitnesses) should recognize and be familiar with specific details of a crime that they committed, whereas innocent individuals would not, thus evoking different ERP patterns.

The majority of ERP studies aimed at detecting deception have involved creatively adapting a Guilty Knowledge Task (GKT) task used in conjunction with the Interrogative Polygraph technique. The GKT (Lykken, 1981) involves asking multiple-choice questions (e.g., what was the murder weapon?). The question is then followed by a series of multiple-choice options, one of which is known to be related to the crime (e.g., the weapon utilized during the crime). The standard interrogative polygraph task assumes that the individual with guilty knowledge should demonstrate increased physiological arousal when the word or picture of the crime-related stimulus is presented (Lykken, 1981).

In addition to adapting the GKT design, ERP studies of deception detection capitalized on the knowledge from the ERP field regarding the extensively researched P300 component. The P300 is a parietally-maximal positive component elicited between 300-1000 ms post stimulus onset. The P300 is reliably elicited to the presentation of rare (i.e., infrequently presented) and meaningful task-relevant stimuli (i.e., targets) embedded within a series of irrelevant stimuli, [a study design that is generically referred to as an odd-ball paradigm (Donchin, 1981; Duncan-Johnson & Donchin, 1979; Johnson, 1986)]. Previous research has also demonstrated that rarely presented, previously learned words embedded within a series of novel words evokes a P300 response (Fabiani, Karis, & Donchin, 1983). This suggests that the P300 can distinguish rarely-occurring recognized from frequently-occurring unrecognized stimuli.

ERP deception detection research has predominantly focused on constructing a GKT within an odd-ball framework (i.e., an odd-ball GKT). The basic assumption behind the ERP approach is that guilty knowledge can be detected by the elicitation of a P300 response to rarely occurring crime-related information (probe) that it is embedded within a series of crime-irrelevant information (irrelevant stimuli). More specifically, the oddball GKT procedure involves presenting a series of options (e.g., six) with one being crime-related and the other five not related. In contrast to individuals with the guilty knowledge, innocent individuals should not demonstrate a P300 response to the crime-related information because they would not be aware of its significance. The waveforms elicited by the irrelevant items can then be compared to the waveforms elicited by the crime-relevant stimuli.

One additional aspect of the oddball GKT task has been the inclusion of an arbitrary target (that is not the crime-related stimuli in question) that demands a unique button press response during the ERP task. Firstly, the button response to the target acts as a control to ensure that participants are responding and carrying out discriminations among the stimuli as well as paying attention to the crime-related stimuli. In the absence of target stimuli, no button presses or unique responses would be required. Secondly, the use of a target can be used as a gauge for an individual's P300 response and therefore provide a baseline for which to compare the probe. The use of irrelevant stimuli provides a baseline for information that is not related to the crime. When a participant is innocent, their ERP responses to crime-related stimuli should be indistinguishable from their response to the irrelevant stimuli. In this manner, innocence or guilt can be determined by the degree of waveform similarity to the irrelevant and the target stimuli. The background literature on ERPs and deception detection will be described in the Introduction section of Experiment 1.

Thesis overview

The proposed research projects aim to investigate the use of ERPs in an eyewitness context, as well as to ascertain the neurophysiological impact of several variables that have been demonstrated to influence eyewitness accuracy.

Experiment 1 assessed whether or not the photograph of the criminal elicited a specific ERP response compared to the photographs of the other lineup members (standard condition). In addition, Experiment 1 also assessed whether a specific ERP response was elicited to the presentation of the criminal when participants actively tried to conceal their knowledge of the criminal (deception condition). For the deception condition, participants were asked to pretend that the criminal was a close friend or relative that they wanted to protect.

Experiment 2a assessed the impact of three different time delays between when participants viewed a crime until when they completed the ERP-lineup (no-delay, 1-hour delay and 1-week delay conditions). Lastly, Experiment 2b examined the neurophysiological impact of whether or not the photograph of the criminal was in the ERP-lineup (CP versus CA condition).

For the set of experiments conducted in this thesis, a crime video was presented followed by a lineup task while participants' ERPs were recorded (i.e., ERP-lineup). Similar to the odd-ball GKT used in the ERP deception detection literature, the ERP lineup consists of a set of six photographs, one of which is the suspect (i.e., the criminal in a CP lineup or an innocent suspect in CA lineups) embedded within a set of five fillers. In addition, an experimenter-imposed target was added to maintain attention and responsiveness to the photographs. This is of particular concern in the deception and CA conditions, as well as in the CP conditions when participants are unable to identify the criminal. For the present set of eyewitness experiments, a photograph of the victim was selected as a target. The victim and criminal share several overlapping features, for example, both are present in the crime video and are newly learned individuals. Therefore, the victim can also act as visual control for the component anticipated to be elicited to the criminal in participants that are able to make a correct identification.

Potential ERP components of interest for the eyewitness ERP-lineup task

The two most likely candidates to be associated with accurate identification of the criminal are 1) the P300, based on the results from the deception detection ERP literature and/or 2) the parietal old/new component, found to be associated with recollection and explicit recognition memory. On one hand, the methodology of the ERP-lineup task for this thesis was designed to mimic the oddball paradigms (i.e., low probability target-detection tasks) that are known to elicit a classic P300 response (Donchin, 1981; Johnson, 1986). Similar to the ERP deception detection studies discussed earlier, the probability of the photograph of the criminal occurring is lower than the probability of the occurrence of photographs of the fillers [e.g., in a six person lineup, the photograph of the criminal occurs 17% (1/6) of the time]. Based on the rarity of the presentation of the criminal, as well as its task-relevance, a P300 response would be predicted. The rare presentation of previously learned items against a backdrop of novel items has also been demonstrated to result in a P300, which distinguishes the recognized from the novel items (Karis, Fabiani, & Donchin, 1984).

However, for the ERP deception detection studies, the Karis et al. (1984) study, the, as well as the ERP-lineup task, the relevant item(s) are selected as meaningful, not by the experimenters, but based on the participant's recognition memory. Therefore, the target and probability effects associated with a P300 response are secondary to or are dependant on the accurate recognition of the crime-related information. For the ERP-lineup task, if participants explicitly recognize and accurately identify the face of the criminal from the preceding crime video (i.e., as an old or previously studied item), then a parietal old/new effect may be the main component of interest. In this manner, one could interpret the study design, not as an oddball paradigm (i.e., one task relevant item that stands out from the others based on its low probability), but as a design in which each face has an equal probability of occurring, one of which is the criminal and thus is more analogous to a recognition memory task. From this perspective, the expected component of interest would be the parietal old/new effect.

Both the P300 and parietal old/new effect components demonstrate temporal and topographical overlap (e.g., both are characterized by a robust positive component elicited within 300-1000 ms range with a parietal maximum). One of the most comprehensive studies designed to segregate these ERP components involved comparing the impact of probability and recognition of old compared to new items (Smith & Guster, 1993). In one condition, old words acted as rare targets (presented 20% of the time compared to 80% new words). In another condition, new words acted as rare targets (presented 20% of the time compared to 80% old words). The results showed that a robust P300 was elicited to rare targets irrespective of whether they were old or new words. However, they also found that old words elicited a parietal old/new effect (referred to by the authors as a memory-evoked shift) compared to new words, irrespective of their probability and whether they were targets or not. These results suggest that both the P300 and parietal old/new effects contribute to the positive deflection within the 300-1000 ms range (Smith & Guster, 1993). Since this study, an extensive body of research has investigated the parietal old/new effect in paradigms independent of probability and target effects (known to elicit a P300) by maintaining an equal probability of the stimulus classes (e.g., 50% old and 50% new items).

It was hypothesized that a positive parietal component will be elicited to correct identification of the criminal and most likely be the result of a both a P300 and a parietal old/new effect. The resulting parietal positive component will subsequently be referred to as a late positive complex (LPC) to encompass the range of positive components that occur within the 300-1000 ms post stimulus onset time window.

From an applied perspective, delineation of the P300 and the parietal old/new effect is considered secondary to the reliable ERP differentiation between the criminal and the fillers. Previous ERP studies for applied or clinical use (e.g., for deception detection or for assessment-based purposes) have highlighted the importance of differentiating between conditions (e.g., guilty versus innocent or correct versus incorrect test responses), rather than on an exploration of the componential features or underlying mechanisms (e.g. Connolly, D'Arcy, Lynn Newman, & Kemps, 2000; D'Arcy, Connolly, & Eskes, 2000; Rosenfeld, 2002).

The current project design involves both the infrequent presentation of the photograph of the criminal, as well as dependence on explicit recognition of the criminal. Therefore the study design does not allow for distinction of the P300 and parietal old/new effect. In line with other ERP research in applied or clinical settings, the main focus of the current set of eyewitness ERP studies is to find reliable components that are able to distinguish accurate identification of the criminal compared to the other lineup members on a group and individual participant level.

Chapter 2: Experiment 1: Standard and Deception Conditions

Experiment 1 Introduction

In the judicial system, decisions regarding an individual's innocence or guilt are guided by physical evidence, expert reports and eyewitness testimony. However, commonly there are controversies regarding the evidence and decisions ultimately rest on the subjective opinions of juries and judges. Over the years, there have been many attempts to develop new techniques and tools to detect when a person is lying or withholding information.

The most widely used instrument aimed at detecting lies in the forensic field has been the interrogative polygraph (commonly referred to as the "lie-detector" test). The premise of this technique is that lying is associated with physiological autonomic nervous system (ANS) activity (e.g., electrodermal responses and heart rate). The rationale is that either the confrontation with one's guilt and/or the fear of detection is believed to elicit an emotional response that can be measured by ANS activity levels. However, there is minimal empirical support for a direct link between lying and ANS activity and this assumption has been extensively challenged on both theoretical and practical grounds (Cross & Saxe, 1992; Kleinmuntz & Szucko, 1982; Kleinmuntz & Szucko, 1984; Rosenfeld, Angell, Johnson, & Qian, 1991; Saxe, 1991; Saxe, 1994; Saxe & Ben-Shakhar, 1999). There is no evidence of a unique pattern of autonomic responses that directly relates to deceptive responding, either within or between individuals (Ben-Shakhar & Furedy, 1990). In essence, the polygraph detects changes in arousal levels, which may or may not be associated with the lie.

Deception detection hit rates using interrogative polygraphy fall slightly above chance levels. However, up to 50% of innocent individuals are incorrectly classified as guilty (Kleinmuntz & Szucko, 1982). Due to the high rates of false positives and negatives, as well as its susceptibility to countermeasures (e.g., coaching strategies to avoid detection), polygraph testing has been considered inadmissible in court in many States in the USA for the last few decades (National Research Council Report, 2003; Saxe & Ben-Shakhar, 1999). Although interrogative polygraphy has had limited success, it is still commonly used by government agencies and private practitioners because its use is believed to lead to an increased number of spontaneous confessions.

Starting in the late 1980's and early 1990's, investigations began to study the detection of concealed information using ERPs. As discussed in the Introduction (Chapter 1), these ERP paradigms were modeled after the Guilty Knowledge Test (GKT) used during interrogative polygraphy (Lykken, 1981). In contrast to techniques using ANS responses as indirect measures of guilt or fear associated with lying, investigators using ERP methods look for the elicitation of ERP patterns associated with the perpetrator's recognition and/or knowledge of crime-related stimuli. Thus, the use of ERPs offers the benefit of an objective measure, based on neurophysiological responses that reflect cognitive processes indicative of knowledge about crime details with which only the perpetrator (or someone directly involved or present during the crime) should be familiar.

Rosenfeld and colleagues conducted many of the early experiments to investigate guilty knowledge detection using ERPs. In an early study, (Rosenfeld, Cantwell et al. 1988), participants were asked to steal one of nine hidden items (e.g., a wallet or a small radio) from a box. For the ERP task, the names of the nine objects were presented multiple times, one of which was the stolen object (guilty group). An innocent control group were also shown nine items, none of which were the stolen item. Only the guilty group were found to elicit a P300 to the presentation of the stolen object.

To increase the ecological validity, Rosenfeld and colleagues conducted another study involving actual misdemeanours committed by participants in the past (Rosenfeld, Angell, Johnson, & Qian, 1991). Subjects were presented with a list of 13 short phrases describing transgressions for which approximately 10-50% of college students were likely to be guilty (e.g., used a false identity card, cheated on a test, smoked pot monthly), but some transgressions were more serious (e.g., stole a car) and less likely (e.g., generally committed by < 2% of college students). Participants were asked to complete a checklist of the transgressions they had previously committed. Participants were misled to believe that the checklist was private (and solely for them to keep track of how many acts they had committed), when in reality participants were monitored by video while they completed the checklist.

Participants were divided into two groups 1) innocent (no reported violations of any of the transgressions) or 2) guilty of at least one but less than five transgressions. One of the 13 transgressions acted as the probe. For the guilty participants, one of their guilty acts was presented as the probe stimulus, whereas for innocent participants one of the transgressions for which they were not guilty of committing acted as the probe. As predicted, a P300 was elicited to the probe only in the guilty participants. Ninety-two percent (12/13) of the guilty and 87% (13/15) of the innocent participants were correctly classified as guilty or not, yielding an overall accuracy rate of 89.3% (Rosenfeld et al. 1991). In addition, the study yielded no false negatives or positives, although a small portion of individuals could not be classified into either of the categories. Furthermore, these results were replicated using a similar paradigm, which yielded comparable identification rates (Johnson & Rosenfeld, 1992). In the follow-up study, participants did not fill out the checklist of transgressions until after the ERP task. This provided further evidence that guilty knowledge from unrehearsed past events (that may have occurred years earlier) could be detected accurately with ERPs.

One study directly compared the use of both the polygraph and ERPs for the detection of guilty knowledge (Farwell & Donchin, 1991). Participants enacted one of two mock espionage scenarios, which involved meeting with someone and exchanging information. Therefore, each participant was guilty for one of the scenarios, but was unaware of the other (and hence "innocent" for the second scenario). The ERP and polygraph were conducted the following day. A P300 was elicited to the espionage-related stimuli only for participants that enacted the corresponding scenario. Eighty-eight percent of the guilty participants were correctly classified. Also, similar to Rosenfeld, Angell et al. (1991), no participants were wrongly classified to innocent or guilty categories. In contrast to the ERP results, the standard polygraph was unable to distinguish guilty from innocent participants above chance levels.

In a second experiment, Farwell and Donchin (1991) investigated the ERP patterns in four individuals with previous criminal experiences. Details about their crimes acted as the probes. All four of the guilty participants were correctly identified by the elicitation of the P300 to the presentation of stimuli associated with their prior offences.

The promising results from these seminal studies led to the growth of ERP research on deception detection (see Rosenfeld, 2002). Throughout the last decade, this field has expanded to include investigation into the detection of malingering of memory deficits and amnesia (Allen, Iacono, & Danielson, 1992; Allen & Iacono, 1997; Ellwanger, Rosenfeld, Sweet, & Bhatt, 1996; Ellwanger, Tenhula, Rosenfeld, & Sweet, 1999; Miller, 1999; Miller & Rosenfeld, 2004; Rosenfeld, et al., 1999; Rosenfeld, Rao, Soskins, & Reinhart Miller, 2003; Rosenfeld, et al., 1998); as well as studies aimed at determining the underlying cognitive progresses associated with deception or the use of strategies to keep track of one's deceptive responses (Johnson, Barnhardt, & Zhu, 2004; Johnson, Barnhardt, & Zhu, 2003; Rosenfeld, et al., 1999).

Other directions in the ERP deception detection literature have attempted to develop, refine and compare statistical procedures with the goal of optimizing guilty knowledge detection rates at the individual level (Allen & Iacono, 1997; Rosenfeld, et al., 1999). Several researchers have attempted to move beyond amplitude and latency characteristics of the P300 component by analyzing scalp distribution patterns to differentiate guilty from innocent profiles based on neural generator differences (Miller, 1999; Miller & Rosenfeld, 2004; Rosenfeld, 2002; Rosenfeld, et al., 1999; Rosenfeld, Rao, Soskins, & Reinhart Miller, 2003).

Overall, deception detection rates using ERPs have yielded approximately 80% accuracy at detecting guilty knowledge, (with a range of about 70% -100%, depending on the paradigm) when using well-rehearsed and/or explicitly remembered information (Allen & Iacono, 1997; Farwell & Donchin, 1991; Girodo, Deck, & Campbell, 2002; Johnson & Rosenfeld, 1992; Miller & Rosenfeld, 2004; Rosenfeld, 2002; Rosenfeld, Angell, Johnson, & Qian, 1991; Rosenfeld, et al., 1999; Rosenfeld, Nasman, Whalen, Cantwell, & Mazzeri, 1987; Rosenfeld, Rao, Soskins, & Reinhart Miller, 2003).

Although, most errors associated with eyewitness testimony were believed to be based on genuine mistaken identity, there have been cases of deliberate perjury (Wells, Lindsay, & Ferguson, 1979; Wells, et al., 1998). There are multiple incentives for perjury, which may include a desire to protect the criminal (e.g., because of fear/threats from the perpetrator, or if they are a family member or close friend), monetary rewards offered by Police Departments for information that will help solve a crime, plea bargains for decreased sentences in exchange for testimony, or the benefits provided to incarcerated informants for information.

For example, in the case of the wrongful conviction of Mr. Harrington *(Harrington, vs. Iowa, 2003)*, the primary eyewitness, Kevin Hughes, changed his testimony on numerous occasions. He sequentially accused three men (each of which turned out to have alibis for the time of the murder) before accusing Harrington. He also changed his testimony regarding the murder weapon on multiple occasions. The Police offered a $5000 reward for information associated with the murder and offered to drop Hughes' prior charges of theft and burglary in exchange for his testimony against Harrington. Of most concern, Hughes' testimony was still permitted even though he had already established himself as a perjurer[1]. In addition, in the cases of wrongful convictions exonerated by DNA evidence described by Conners et al. (1996), 5/28 were associated with deliberate perjury of at least one of the key witnesses, two of which were eyewitnesses (Connors, Lundregan, Miller, & McEwen, 1996).

The current experiment was developed to assess the use of ERPs in an eyewitness context (standard condition) as well as for the use of detecting eyewitnesses who

---

[1]See http://www.judicial.state.ia.us/supreme/opinions/20030226/01-0653.asp?search=01%2d0653#_1

purposely try to conceal their knowledge of the criminal (deception condition). It was hypothesized that, if participants are able to accurately recognize the criminal, a LPC will be elicited to the photograph of the criminal, irrespective of whether the participant tried to conceal their knowledge of the criminal or not. For the deception condition, participants will be provided with a concealing strategy - to identify the criminal as innocent. In a non-ERP eyewitness study, participants were asked to lie to protect the identity of the criminal. It was found that almost 100% of participants elected to reject the lineup, rather than falsely identifying one of the fillers as the criminal (Parliament & Yarmey, 2002). Therefore, the concealing strategy provided for the deception condition most likely reflects the most common strategy used by eyewitnesses who attempt to protect the criminal.

Experiment 1 Methods

Participants

Twenty-two undergraduate students (12 females) with a mean age of 26.1 years (*SD*= 4.9) were recruited from a departmental subject pool and participated in the study for course credit or $5.00/hr. Participants were fluent English speakers with normal hearing and normal or corrected-to-normal vision and no reported history of neurological or psychiatric conditions. The data from two participants (one male and one female) were excluded from the study because of EEG recording difficulties that resulted in loss of data (S07) or low trial numbers following electrooculogram (EOG) artefact rejection (S13), refer to the Electrophysiological recording procedure section below. Therefore, the data for this study was based on 20 participants. Informed consent was obtained at the beginning of the experiment and participants were debriefed at the end of the session. The Research Ethics Board of Dalhousie University approved the study.

Stimuli

Crime videos

Two 1-minute simulated non-violent crime scenarios (videos A and B) were used for this study. For video A, a female walked into an office and placed her purse on a desk. She then went into an adjacent room. A man walked in and rummaged through her purse and stole money from it. The criminal was present for approximately 15 seconds and there were opportunities to view both frontal and profile views. The woman returned to see the man run off with the money. Video B consisted of the same crime scenario with a different criminal.

ERP-lineups:

Two ERP-lineups (A and B) were used for this study that corresponded to the two crime videos (A and B). ERP-lineup A consisted of seven digitized photographs. The photographs consisted of five fillers (F1, F2, F3, F4 and F5), the criminal (C) from video A and the victim (V). The photographs of the criminal and fillers were of males in their 20's, from the waist up looking forward. ERP-lineup B also consisted of seven digitized photographs (F1, F2, F3, F4 and F5, C from video B and V). The fillers were of different individuals for ERP-lineups A and B. For each ERP-lineup, the fillers were selected prior to the study based on both shared attributes (e.g., approximate age, race, hair length) with the criminal as well as based on commonalities from written descriptions of the criminal from a large sample of people after viewing the crime scenarios[2]. The written

---

[2]Filler selection was conducted prior to the current study by Steven Smith.

descriptions consisted of an open-ended paragraph describing the criminal followed by completion of 11 closed-ended questions regarding the age, weight, height, body type, hair colour, length and type, race, facial hair, glasses and eye colour of the criminal. For each ERP-lineup, five photographs of individuals matching the descriptions of the criminal were selected from a database of photographs and acted as the fillers in the lineup. In addition, a mock witness procedure (Doob & Kirshenbaum, 1973) had previously been conducted to rule out lineup bias for both ERP-lineups A and B[3]. In brief, *mock witnesses* were provided with written descriptions of the criminal and were asked to try and identify the criminal from the lineup without ever actually seeing the crime video. A lineup was considered biased if the criminal was reliably identified. The two lineups (A and B) selected for this study were considered non-biased based on this procedure prior to commencement of the experiment.

Procedure

Participants were seated in a comfortable chair for the duration of the experiment. The electrode set up commenced immediately (details of the electrophysiological recordings and procedures are described below in the Electrophysiological recording procedures section). Following the electrode set-up, participants watched one of the two non-violent simulated crime videos (the order of the videos were counterbalanced across participants). Prior to watching the first video, participants were instructed to pay attention and informed that they will later be asked to try to identify the criminal from a lineup. Upon completion of the crime video, participants were asked to pretend that the video they just watched was a real crime for which they were the only eyewitness. They

---

[3]The mock witness procedure was conducted prior to the current study by Steven Smith.

were also asked to pretend that they were at a police station and being asked to try to identify the criminal from an ERP-lineup.

Approximately one minute following the end of the first video (e.g., video A), participants were presented with the ERP-lineup that corresponded to the video presented (e.g., ERP-lineup A) while their EEG was recorded. The photographs were presented and formatted to fit on a 21-inch computer monitor, which was placed 1.5 m in front of participants at eye level. Each of the seven photographs were pseudorandomly presented 40 times (i.e., all seven photographs were randomly presented prior to the repetition of any lineup member) for a duration of 1500 ms with an inter-stimulus interval of 1500 ms.

Following the presentation of each photograph, participants were asked to make a decision about the identify of each photograph by pressing one of three buttons on a key pad, which corresponded to a decision that the photograph was the i) the criminal ii) the victim or iii) a filler (standard condition). Specifically, participants were asked to *select the criminal* from the series of photographs based on their button press responses. Participants were encouraged to respond accurately rather than quickly and to make their decision about the identity of each lineup member following the first presentation of each photograph. The ERP-lineup took approximately 14 minutes to complete. Following completion of the first ERP-lineup, participants were sequentially shown the photographs of the lineup members and asked to rate each individual on a scale of 0-10 (referred to as *certainty ratings).* A certainty rating of 10 indicated that they were certain that the individual in the photograph was the criminal and a 0 indicated that they were certain that the individual was not the criminal.

Participants then watched the second crime video (e.g., video B) and completed the second corresponding ERP-lineup (e.g., ERP-lineup B). The second ERP-lineup followed the same procedure, however, participants were asked to pretend that the criminal was a close friend or relative (or someone whom they would want to protect). They were then asked to try to conceal the identity of the criminal from the researchers by classifying the criminal as innocent (i.e., as a filler) during the button press component of the ERP-lineup task (deception condition). Following the second ERP-lineup, participants were asked to provide certainty ratings for each of the photographs on a scale of 0 to 10. Participants were asked to provide truthful certainty ratings for the deception condition in order to evaluate their identification decisions. The standard condition was always presented before the deception condition as a precaution to decrease the likelihood of participants adopting a deceptive strategy during the standard condition.

Electrophysiological recording procedures

EEG was recorded using an electrode cap with Ag/AgCl electrodes that was placed on participants' head. The cap contained 30 electrode scalp sites in accordance with the 10-20 electrode system (Jasper, 1958) referenced to linked earlobes: Fp1, Fp2, F7, F3, Fz, F4, F8, Ft7, Fc3, Fcz, Fc4, Ft8, T7, C3, Cz, C4, T8, Tp7, Cp3, Cpz, Cp4, Tp8, P7, P3, Pz, P4, P8, O1, Oz and O2. Participants were grounded with a forehead electrode. The EOG was recorded from electrodes placed above and below the right eye (vertical, VEOG) and from electrodes lateral to each eye (horizontal, HEOG). Prior to placement of the EOG and earlobe electrodes, the areas were cleaned with an exfoliating cream (Neuroprep©). Electrolyte gel was placed between the electrodes and the scalp or skin for electrode conductance. Electrode impedances were kept below 5 kΩ.

The EEG was amplified 10,000 times and recorded with a bandpass of 0.01-100 Hz and digitally sampled at 500 Hz. Event-related brain potentials (ERP) recordings were time-locked to the onset of the presentation of the seven photographs (F1, F2, F3, F4, F5, C and V) and epoched for 800 ms (including a 100 ms pre-stimulus baseline). Each photograph was presented and recorded 40 times. The epoched data was filtered off-line with a bandpass of 0.1-20 Hz (24 dB/octave). Trials with EOG voltages greater than +/- 75 μV were discarded from the analyses because they mask the ERP effects of interest. Following EOG artefact rejection, a mean of 85.4% (*SD* = 10.9, range = 63.6 – 97.5%) of the data was retained for the analyses in the standard condition and 86.6% (*SD* = 8.2, range = 66.7 - 99.6%) for the deception condition.

The ERPs from the EOG artifact-free data for each individual participant were averaged according to photograph (F1, F2, F3, F4, F5 and C) for both the standard and deception conditions. In other words, the 40 recordings of each photograph (minus the discarded trials with EOG artifacts) were averaged to create six waveforms for each individual for each condition. Grand (or group) average waveforms were created for each photograph (F1, F2, F3, F4, F5 and C) for each condition by averaging the individual averages.

Experiment 1 Results

Grand average waveforms and ANOVA analyses

Figures 1 and 2 depict the grand average waveforms elicited to the criminal and the fillers combined (i.e., the waveform that results from averaging the grand averages of each of the fillers) across the 30 scalp electrodes for the standard and deception conditions, respectively. Of most relevance, a visually apparent LPC was elicited to the photograph of the criminal (maximal at the centro-parietal electrode sites) for both conditions in comparison to the grand average waveform of the fillers combined.

Figures 3 and 4 depict a more detailed view of the grand average waveforms at the six centro-parietal electrode sites (Cp3, Cpz, Cp4, P3, Pz and P4) for each lineup member in the standard and deception conditions, respectively. Based on these figures, a robust LPC was visually apparent at all six of the electrode sites to the criminal compared to each of the fillers between 400-600 ms post-photograph onset. Therefore, the mean amplitude of this time window (referred to as the *fixed-interval* analysis) was selected as the main area of interest for the six scalp electrodes (Cp3, Cpz, Cp4, P3, Pz and P4).

In addition, a similar set of analyses was conducted based on *a priori* knowledge of the criminal (referred to as the *criminal-based* analysis). The fixed-interval analysis has the advantage of remaining non-biased toward any of the lineup members, whereas the criminal-based analysis uses the *a priori* knowledge of the lineup member suspected to be the criminal. One-suspect lineups are highly recommended over multi-suspect lineups (e.g. Wells, et al., 1998) because the likelihood of false identifications is greatly increased as the number of suspects increases (Wells & Turtle, 1986). Thus, if a lineup contains only one suspect of interest, then isolating and tailoring the analysis to that suspect may be of benefit. The criminal-based analysis also has the benefit of tailoring the analysis to the 100 ms latency range of most interest for each individual. For the fixed-interval analysis the dependant variable was the mean amplitude of the 400-600 ms latency window. In contrast, for the criminal-based analysis, the peak of each individual's average waveform to the photograph of the criminal was first identified and then the mean amplitude +/- 50 ms surrounding the peak of the criminal was selected as the dependant variable. The mean amplitude for this latency range was then calculated and compared for each of the lineup members.

Fixed-interval analyses

A 3-way repeated measures analysis of variance (RM ANOVA) with CONDITION (standard and deception), PHOTO (F1, F2, F3, F4, F5 and C)[4] and SITE (Cp3, Cpz, Cp4, P3, Pz and P4) as factors was conducted (refer to Table 1) and subjected to a Greenhouse-Geisser conservative degrees of freedom correction (Greenhouse & Geisser, 1959). Only significant main and interaction effects considered germane to the experimental hypotheses were explored (i.e., significant effects collapsed across PHOTO were not explored). For both the fixed-interval and criminal-based analyses, planned comparisons were computed comparing the criminal to each filler (for both conditions). Other relevant significant effects were further analyzed by post-hoc comparisons with a Bonferroni correction.

The significant main effect of PHOTO and subsequent planned comparisons reflected an increased LPC to the criminal compared to each of the fillers. The interaction effects are best interpreted within the significant CONDITION x SITE x PHOTO interaction. Therefore, post hoc tests were computed only on this interaction effect. In line with the findings from the PHOTO main effect, the LPC was significantly larger to the criminal compared to each filler at all the six electrode sites for both the standard and deception conditions. Comparing across conditions, the LPC was significantly greater to

---

[4]The photograph of the victim was not included in the analyses. The victim was included in the task design to maintain attention and responsiveness to the photographs. The victim was also used as a visual control for the LPC to the criminal. Generally, the waveform to the victim elicited an LPC component similar to the criminal in the standard condition compared to the remaining lineup members.

the criminal in the standard than in the deception condition at each of the six electrodes. This finding was expected based on past ERP deception detection studies that demonstrated diminished amplitude to the crime-related probe compared to target stimuli when participants attempted to conceal the information (see Rosenfeld, 2002).

When comparing across sites, the LPC elicited to the criminal in the standard condition demonstrated a midline-right centro-parietal scalp voltage distribution (Pz, Cpz) > P4 > P3 > (Cp4, Cp3). A similar scalp voltage distribution pattern was seen for the deception condition (Pz, Cpz, P4) > (P3, Cp4) > Cp3).

Criminal-based analyses

The RM ANOVA results from the criminal-based analyses (Table 1) yielded almost exactly the same results as the fixed-interval analysis and therefore will not be discussed.

Individual ERP analyses

Figures 5 and 6 depict the waveforms elicited to the criminal and the fillers combined for each individual participant at the Pz electrode site for the standard and deception conditions, respectively. For each participant, the mean amplitude between 400-600 ms was calculated from the artifact-free single sweeps for each of the six lineup members (F1, F2, F3, F4, F5 and C). A procedure analogous to a deviation contrast set analysis was conducted. For this procedure, the mean (*M*) and standard deviation (*SD*) of five (e.g., F1, F2, F3, F4 and F5) out of the six lineup members (F1, F2, F3, F4, F5 and C) were calculated. The mean amplitude of the remaining photograph (e.g., C) was then compared to determine if it fell beyond two standard deviations (i.e., ≥ a z-score of 2) from the distribution. This procedure was conducted for the six centro-parietal scalp sites

(Cp3, Cpz, Cp4, P3, Pz, P4) and repeated five times, once for each lineup member (e.g., the mean and standard deviation was calculated for F1, F2, F3, F4 and C and then a z-score was computed indicating how far the mean amplitude of F5 deviated from the distribution of means). Then, the z-scores associated with each photograph across the six scalp sites were averaged (mean z-score).

The mean z-score provides an index that designates which of the lineup members, if any, are significantly differentiated from the others. It was hypothesized that if participants were able to make a correct identification, then the mean z-score associated with the criminal should be largest and differentiated from each of the fillers. The same procedure was also conducted based on the criminal-based latency window for each participant.

A *correct identification* was assigned if the highest mean z-score was associated with the photograph of the criminal and it exceeded a cut-off z-score ≥ 2. A *misidentification* was assigned if the highest mean z-score was associated with one of the fillers and was ≥ 2. No *false identifications* were possible because the remaining lineup members were known innocent fillers. Lastly, a *false rejection* (i.e., belief that the criminal was absent from the lineup) was assigned if the mean z-score associated with any of the lineup members failed to be ≥ 2. The same procedure was applied to the data from the criminal-based analysis. The percentage of correct identifications, misidentifications and false rejections were calculated for both conditions.

Tables 2 and 3 show the individual ERP z-score results for ERP-lineups A & B for the standard condition, for the fixed-interval and criminal-based analyses, respectively. Tables 4 and 5 show the individual ERP mean z-score results for ERP-lineups A & B for the deception condition, for the fixed-interval and criminal-based analyses, respectively. Although on a group level the fixed-interval and criminal-based analyses did not differ, there were a few differences on an individual level. Therefore, the results from both analyses will be described.

Standard condition

In the standard condition, the mean z-score was found to be highest to the criminal in 95% (19/20) of participants, where 85% (17/20) had a mean z-score ≥ 2 for both the fixed-interval and criminal-based analyses. In this condition, one participant (S01) demonstrated the highest mean z-score to a filler (based on both the fixed-interval and criminal-based analyses). However the mean z-score to the filler was ≤ 2, indicating a false rejection. Overall, the correct identification rate was 85% (17/20), and false rejections were 15% (3/20) for both the fixed-interval and the criminal-based analyses. There were no misidentifications.

Deception condition

For the fixed-interval analyses, the mean z-score was highest to the criminal in 65% (13/20) of participants, 50% (10/20) had a mean z-score ≥ 2. For the criminal-based analysis 75% (15/20) of participants exhibited the highest mean z-score to the criminal, and 55% (11/20) had a mean z-score ≥ 2. In the remainder of the participants, the LPC patterns reflected 20% (4/20) misidentifications and 30% (6/20) false rejections, based on the fixed-interval analysis. For the criminal-based analysis, 5% (1/20) made a misidentification and 40% (8/20) of participants made false rejections. Thus, the LPC patterns were able to distinguish correct identifications in the majority of participants that attempted to conceal their knowledge of the criminal.

Criminal A versus Criminal B

It was speculated that there may be increased LPC amplitude for criminals with distinctive attributes that make them easier to identify relative to faces with more typical characteristics. Based on reports from participants, Criminal B was described as more distinct and easier to identify than Criminal A. Although the sample size was too small to conduct statistical analyses, from Figure 7, there was a visually-apparent larger LPC to Criminal B compared to Criminal A for the standard and deception conditions. Interestingly, the LPC differences between Criminals A and B predominantly occurred within the 300-400 ms range for the deception condition.

Additionally, for the standard condition, ERP-lineup A, 90% (9/10) of participants demonstrated the highest mean z-score to the criminal for both the fixed-interval and criminal-based interval analyses. Seventy percent (7/10) had a mean z-score ≥ 2. For ERP-lineup B, the highest mean z-score was associated with the criminal in 100% (10/10) of participants, all of which had a mean z-score ≥ 2, for both the fixed-interval and criminal-based analyses. For the deception condition, for ERP-lineup A, the mean z-score was highest to the criminal in 60% (6/10) of participants, 50% (5/10) with a mean z-score ≥ 2, for the fixed-interval analysis. For the criminal-based analysis 70% (7/10) demonstrated the highest mean z-score to the criminal, 50% (5/10) of which were ≥ 2. For ERP-lineup B, the highest mean z-score was to the criminal in 70% (7/10) of participants (50% or 5/10 with a mean z-score ≥ 2) for the fixed-interval analysis, and 80% (8/10) of participants (60% or 6/10 with a mean z-score ≥ 2) for the criminal-based analysis. These results also demonstrate that Criminal B was associated with more correct identifications based on the mean z-scores.

Identification accuracy

Certainty ratings

Certainty rating were based on a 0-10 scale, where a 10 indicated that the participant was certain the photograph was of the criminal and a 0 indicated that they were certain that the photograph was of a filler and a 5 indicated that they were unsure if the photograph was of the criminal or not. Eyewitness's accuracy for each condition was determined based on the percentage of participants that selected the highest certainty rating for the photograph of the criminal. Although participants in the deception condition were asked to conceal their knowledge of the criminal during the ERP-lineup, they were asked to provide truthful certainty ratings. A *correct identification* was classified when participants' highest certainty rating was to the criminal and was ≥ 5, otherwise it was counted as a rejection of the lineup (i.e., an indication that the participant believed that the criminal was absent from the lineup). In the case where a participant provided the same highest rating for more than one lineup member, they were asked to indicate which one they felt was the criminal, if any. A *misidentification* was classified when participants rated one of the fillers with the highest certainty rating that was above 5. A *false rejection* was counted when the certainty ratings to all the lineup members were below 5 in a CP lineup. The percentage of correct identifications, misidentifications and false rejections were calculated for both conditions.

Tables 6 and 7 depict participants' certainty ratings for each of the photographs for both the standard and deception conditions, respectively. All participants (except S11 in the standard condition and S03 in the deception condition) rated the criminal with the highest certainty rating. S11 rated two lineup members with a certainty rating of 5 in the standard condition and S03 rated three of the lineup members (including the criminal) with a certainty rating of 5 in the deception condition. When the highest certainty rating was selected for more than one lineup member, participants were asked which one they felt was the criminal. Both S11 and S03 selected the criminal in response to this question. This response was also reflected by their button press responses. Overall, correct identification rates were 100%. There were no false rejections or misidentifications for both the standard and deception conditions.

Button press accuracy

Eyewitness' accuracy was determined by calculating the percentage of time participants classified each lineup member as the criminal. Participants' responses were classified as a correct identification, a misidentification or a false rejection. A *correct identification* was classified if participants correctly selected the criminal most frequently and a least 50% of the time. A *misidentification* was classified if participants implicated one of the fillers as the criminal most frequently and at least 50% of the time. A *false rejection* was classified when participants failed to consistently identify any lineup member as the criminal (i.e., they did not select any lineup member as the criminal more than 50% of the time). The number and percentage of correct identifications, misidentifications and false rejections were calculated.

Tables 8 and 9 show the percentage of times participants' classified each lineup member as the criminal based on their button press responses, in the standard and deception conditions, respectively. For the standard condition, all participants correctly classified the criminal most frequently and at least 50% of the time. In the deception condition, all participants concealed their knowledge of the criminal by classifying the criminal as a filler 100% of the time.

Comparison of identification accuracy measures

Tables 10 and 11, provide a summary chart for the number and percentage of correct identifications, misidentifications and false rejections based on certainty ratings, button press responses and ERP mean z-scores from the criminal-based analysis, for the standard and deception conditions, respectively.

Experiment 1 Discussion

The primary goal of the present study was to investigate the use of ERPs in an eyewitness lineup identification task. As a secondary interest, the present study aimed to assess the impact of deception (or concealment of the criminal) on the ERP patterns. In line with the hypotheses, the major finding was that a centro-parietally-based LPC (maximal between 400 and 600 ms post-photograph onset) was elicited to the criminal compared to each of the fillers, both when participants were trying to accurately identify the criminal (standard condition) and when they attempted to conceal their recognition of the criminal (deception condition). This pattern was evident both collapsed across electrode sites as well as at each of the six centro-parietal sites individually.

The strong pattern of results on a group level is highly encouraging, but it is essential to demonstrate reliable patterns on an individual level in order for the results to be applicable to an applied eyewitness setting. Two types of analyses were conducted for the individual participants, one method was based on the mean amplitude of the 400-600 ms window (fixed-interval analysis) and the other was based on the mean amplitude surrounding +/- 50 ms of the peak of the criminal (criminal-based analysis). The fixed-interval analysis has the advantage of remaining non-biased toward any of the lineup members, whereas the criminal-based analysis uses the *a priori* knowledge of the lineup member suspected to be the criminal. Additionally, the criminal-based approach allows for the analysis to be targeted to the 100 ms latency window of most interest for each individual and thus decreases the individual variability that may dampen effects in the fixed-interval approach.

The criminal-based analysis yielded the same or higher rates of correct identifications compared to the fixed-interval analysis based on the ERP differentiation of the criminal from the fillers. Therefore, for simplicity, only the results from the criminal-based analyses will be discussed. For the standard condition, the criminal was distinguished from each of the fillers in 95% of participants (85% of which exceeded the statistical mean z-score ≥ 2 cut-off to be classified as a correct identification). In addition, the remainder of the participants made false rejections rather than misidentifications. Thus, there was no risk of wrongful accusations. In the deception condition, the criminal was differentiated from the fillers in 75% of participants, but only 55% exceeded the statistical cut-off.

Impressively, differentiation of the criminal from the fillers was achieved in 100% of participants based on the highest z-score at the Cpz and Pz electrode sites in the standard condition (80% of which exceeded the statistical cut-off). This finding is important because previous ERP studies of deception detection have focused their analyses on the Pz site (Allen, Iacono, & Danielson, 1992; Farwell & Donchin, 1991; Rosenfeld, Angell, Johnson, & Qian, 1991). The results provide replication of the significance of this scalp region. In addition, from an applied perspective, a smaller electrode array would mean less set-up time and simpler analyses and thus increase the likelihood of use in real-world settings.

Different rates of correct identifications based on the mean z-scores were found for ERP-lineup A and B. In the standard condition, Criminal A was distinguished from the fillers in 90% of participants, 70% of which exceeded the statistical the cut-off, compared to 100% participants, all of which reached the cut-off, for Criminal B. Similarly, in the deception condition, Criminal A was distinguished from the fillers in 70% of participants (50% exceeded the statistical cut-off) compared to 80% of participants (60% exceeded the statistical cut-off) for Criminal B. Identification accuracy based on certainty ratings also demonstrated that participants were more certain about the identity of Criminal B compared to Criminal A. For the standard condition, 90% of participants rated Criminal B with a certainty rating of 10, indicating that they were certain that the photograph was of the criminal, compared to only 60% of participants for criminal A. Similarly, in the deception condition, 90% of participants rated Criminal B a 10, compared to only 40% for Criminal A.

Although the features of what makes a face more memorable than another can be difficult to objectively define, research has consistently supported that distinct faces are more memorable (e.g. Going & Read, 1974; Light, Kayra-Stuart, & Hollander, 1979; Wickham & Morris, 2003). Distinctness is commonly measured by either deviations from proportions from average facial dimensions or by asking participants to rate faces based on the *ease of spotting the face in a crowd.* Criminal A, was a Caucasian male of average height, with short brown hair, an average build, and no distinct facial features. In contrast, Criminal B was a Caucasian male with a thin build, above average height, long hair and narrow facial features. Although distinctness measures were not collected for the criminals used in this study, it is reasonable to propose that Criminal B was more distinct. It was believed that the greater distinctness of Criminal B led to easier recognition, which was reflected by the trend of increased LPC amplitude and higher certainty ratings. An alternative explanation is that Criminal B may have been easier to identify as a result of a biased lineup or difficulty finding fillers with overlapping features compared to Criminal A. However, both of the lineups were found to be non-biased prior to the initiation of the study. Therefore, we believe that increased LPC patterns are associated with stronger recognition that results from the increased distinctness of Criminal B.

The rates of detecting correct identifications on an individual level based on ERP patterns were higher for the standard compared to the deception condition. In the eyewitness field, it is believed that the majority of false identifications are most likely a result of well-intentioned individuals who are genuinely mistaken, rather than cases of deliberate perjury (e.g. Wells, Lindsay, & Ferguson, 1979; Wells, et al., 1998). Thus, although cases of deliberate perjury do exist, it is not the major factor leading to the wrongful convictions (Connors, Lundregan, Miller, & McEwen, 1996; Wells, Lindsay, & Ferguson, 1979; Wells, et al., 1998). However, the lack of objective verification of eyewitness reports makes detecting perjury very difficult. Therefore, the optimistic results of this study provide evidence that the use of ERPs in an eyewitness context may aid in determining the accuracy of eyewitness identifications.

Although the LPC emitted to the criminal was decreased in the deception compared to the standard condition, the ERP patterns were still able to detect the majority of participants that concealed their knowledge of the criminal's identity. ERP deception detection studies generally report an average of 80% correct detection of guilty knowledge (see Rosenfeld, 2002). Therefore, the rates of deception detection using ERPs were higher in other studies compared to the current study. This may be related to the type of individual statistical analyses conducted on our data. Unlike the previous ERP deception detection studies that group all stimuli with the exception of the crime-related probe and the arbitrarily selected target into one class of stimuli, the current study compared the criminal to *each* of the fillers separately. This type of analysis was believed to be important for application in an eyewitness context. Although the criminal was the only suspect in the lineup, it was important to segregate each of the fillers to ensure that only one lineup member stands out compared to the remaining members. There is a potential risk of masking an LPC response to one of the fillers due to the signal averaging process of combining all fillers together. Thus, our individual analyses are much more stringent but also, in our view, more valid.

In addition, most previous ERP deception detection studies used a bootstrapping procedure [adapted from (Wasserman & Bockenholt, 1989)] to increase the statistical power of the ERP data on an individual level (Allan & Rugg, 1997; Allen, Iacono, & Danielson, 1992; Farwell & Donchin, 1991; Rosenfeld, Angell, Johnson, & Qian, 1991). Bootstrapping acts to provide an estimate of the actual distribution of a given parameter when there are only a limited number of samples, which is often the case for ERP data on an individual level. For bootstrapping, a series of observations are randomly selected (with replacement) and averaged together (subsamples). This is done multiple times. Then, the subsamples are averaged together. Thus, bootstrapping simulates multiple data sets that are presumed to reflect the actual distribution. Percentile confidence intervals can then be produced.

The use of this technique is beneficial if trial numbers or effect sizes are too low to obtain a significant differentiation between conditions or if there is speculation that the data is not normally distributed. However, for a bootstrapping technique, the same observations are used multiple times, creating a regression to the mean, biasing the data and violating assumptions about random sampling distributions. A bootstrapping approach was not applied to the data in the present study because the differentiation of the criminal appeared to be strong enough to demonstrate significant differences without undergoing this form of statistical manipulation. However, it is possible that some cases where the criminal could not be statistically distinguished from the fillers may be masked due to a low SNR or lack of power. This may be particularly the case in the deception condition where an LPC amplitude difference (or effect size) between the criminal and fillers was less pronounced.

Although participants were asked to conceal their knowledge of the criminal during the ERP-lineup, they were asked to provide truthful certainty ratings upon completion of the ERP-lineup task. All participants were able to accurately identify the criminal in both the standard and deception conditions based on their certainty ratings. However, one participant in the standard condition and one participant in the deception condition rated one of the fillers with an equally high certainty rating. These participants were asked to select only one member of the lineup as the criminal, and in both cases the participants selected the criminal. In terms of button press responses, in the standard condition, all participants correctly classified the criminal from the ERP-lineups with over 90% accuracy (with the exception of one participant, S17, who correctly classified the criminal 80% of the time). In the deception condition, all participants successfully concealed their knowledge of the criminal by classifying the criminal as innocent as determined by their button press responses.

It is interesting to note the high degree of identification accuracy based on certainty ratings and button press accuracy (100%) in this study in contrast to previous behavioural studies and meta-analyses that have used similar methodologies. Based on three meta-analyses, correct identifications for CP lineups were found to be 51% (Haber & Haber, 2001), 45% (Steblay, Dysart, Fulero, & Lindsay, 2003) and 35% (Steblay, Dysart, Fulero, & Lindsay, 2001). Misidentifications were 27%, 24% and 19%, and false rejections were 21%, 34% and 46%, respectively for the three meta-analyses[5].

Although it is unknown why the behavioural accuracy rates from the current study were higher than past studies using similar methodologies, it is speculated that the ERP-lineup procedure may have facilitated accuracy. The repetitive viewing of each photograph 40 times may have increased participants' degree of certainty and confidence in their identification decisions. However, because none of the participants made misidentifications, it is unknown if the ERP procedure would have also inflated participants' certainty in their decision if they made misidentifications. In addition, the use of biased lineup instructions (e.g., select the criminal from the lineup) may have also inflated the accuracy rates. It has previously been shown that use of this type of lineup instruction (although common in real-world situations) motivates participants to rely on a

[5]Only the results from sequential lineups are reported for Steblay, Dysart, Fulero, & Lindsay (2001), whereas the other meta-analyses results are collapsed across simultaneous and sequential lineup presentations.

relative judgment strategy and to guess under the circumstances where they would otherwise be unsure (Clark, 2005; Wells, et al., 1998). Past studies have demonstrated that guessing in these contexts increases correct identifications in CP lineups compared to studies that use non-biased instructions [e.g., *the criminal might or might not be present in the lineup* (Clark, 2005; Wells, et al., 1998)].

For this study, a deception strategy was imposed - participants were asked to conceal their knowledge of the criminal by classifying the criminal as an innocent lineup member. This has previously been demonstrated to be the dominant strategy used by participants when they were asked to lie to protect the perpetrator [e.g., participants tend to reject the lineup rather than make deliberate misidentifications (Parliament & Yarmey, 2002)]. However, it is possible that some actual eyewitnesses may combine the strategy imposed in this study (i.e., choosing not to identify the criminal) in combination with intentionally selecting a filler as the criminal. The impact of this tactic on ERP patterns is unknown. However, the reliable selection of one of the fillers may lead to the elicitation of a classic P300 response due to the resulting *targetness* effect that may develop to the selected filler. Thus, this type of combined strategy may make it challenging to differentiate the waveforms from the actual criminal compared to a falsely accused lineup member.

In line with this speculation, one study investigated the use of countermeasures during two ERP-GKT tasks (Rosenfeld, Soskins, Bosh, & Ryan, 2004). The countermeasure strategy was to generate covert responses to the presentation of all the task stimuli (e.g. wiggle your left toe), with the goal of increasing the meaningfulness and task relevance of the irrelevant stimuli, and hence reducing a P300 odd-ball effect. When the countermeasure strategy was used, the ability to detect deception dropped from 82% to 18% for a six-probe ERP-GKT [adapted from the task developed by Farwell and Donchin (1991)]and from 92% to 50% for a one-probe ERP-GKT [adapted from the task developed by Rosenfeld et al. (1991)]. Although this study does not investigate the impact of the use of a combined strategy of concealing the identity of the criminal, while purposely misidentifying a filler, it does suggest that an overt response to a filler would impact the ERP results and make the concealment of the criminal harder to identify. More research is needed to investigate the impact of such a strategy on ERP patterns.

Generally, it is recognized that eyewitness identification is prone to errors and biases (Wells, 2001; Wells & Loftus, 2003; Wells & Olsen, 2003; Wells & Seelau, 1995; Wells, et al., 1998; Wells & Turtle, 1987). However, eyewitness reports offer an essential piece of evidence in real-world cases (Wells & Loftus, 2002; Wells, et al., 2000). The problem is that there is no clear-cut way of determining which eyewitnesses are accurate in their identifications and which are not. In fact, the commonly believed notion - that an eyewitnesses' confidence in their identification decision is linked with accuracy, has failed to demonstrate a high correlation (Cutler, Berman, Penrod, & Fisher, 1994; Cutler & Penrod, 1989a; Cutler & Penrod, 1989b; Cutler, Penrod, & Dexter, 1990). Lineup decisions have important and serious social and legal consequences and therefore an objective measure would be of great benefit. The group and individual results from the present study lend encouraging support for the ability of ERPs to provide a neurophysiological index to help provide an objective measure of eyewitness accuracy, even when participants attempt to conceal their knowledge about the criminal.

Chapter 3: Experiment 2a: No-Delay, 1-Hour Delay and 1-Week Delay Conditions

Experiment 2a Introduction

A major issue in the eyewitness field is the impact of the time delay between witnessing the crime and viewing a lineup. In real world criminal cases, this time interval is highly variable and, in some cases, can be up to years. There appears to be unanimous support for the notion that identification accuracy rates decrease as time passes.

Mathematical models of forgetting date back to Ebbinghaus (1885) and have suggested that decay of recall memory follows a simple power function over time, which is characterized by an initial rapid decrease, followed by a long and slow decay. This power function has been found to be applicable to multiple aspects of memory. For example, this pattern has been reliably reproduced in humans, across multiple stimulus types (e.g., for a series of memorized words, objects or pictures) and retention intervals (e.g., ranging from a few seconds to weeks). This pattern has also been seen in delayed match-to-sample memory tasks in animals (Wixted & Ebbesen, 1991; Wixted & Ebbesen, 1997).

Despite the ubiquitous nature of a power function decline of memory, the scale of the decay can change dramatically depending on the impact of multiple variables. Research comparing recognition across various stimulus classes (e.g., words vs. pictures vs. faces) has demonstrated that memory for pictures is superior and less susceptible to decay than verbal material (Haber, 1970; Shepard, 1967; Standing, 1973) and superior for human faces compared to pictures of buildings (Scapinello & Yarmey, 1970). One study (Deffenbacher, Carr, & Leu, 1981) found that memory for both nouns and pictures of objects was superior to faces at short retention intervals (2-minute), but that memory was better preserved for faces over longer retention intervals (2-week). These different recognition decay rates across stimulus classes have been interpreted as evidence for separate memory systems for verbal, object, and face stimuli (Deffenbacher, Carr, & Leu, 1981; Paivio, 1969). Overall, research suggests that the memory decay curve is steepest for verbal stimuli followed by pictorial and face stimuli, respectively.

Studies specifically testing the impact of time delay for briefly presented (i.e., a few seconds) pictorial stimuli have demonstrated that recognition memory is generally unaffected for delays of up to five hours but then drops off as the time delay increases (e.g., from one day to a year), although a substantial amount of retention remains at the longer time delays (Arnoult, 1956; Nickerson, 1968; Park, Royal, Dudley, & Morrell, 1988; Shepard, 1967). Moreover, recognition accuracy of faces previously presented only once can generally be sustained for up to 30 days without statistically significant decay (Chance, Goldstein, & McBride, 1975; Davies, Ellis, & Shepherd, 1978; Shepherd & Ellis, 1973).

Of further interest, increased familiarity of stimuli (e.g., as a result of multiple repetitions, longer time to study or more contact with the stimuli) has consistently been shown to be associated with higher retention maintenance over time (Arnoult, 1956; Nickerson, 1968; Scapinello & Yarmey, 1970; Shepard, 1967). In fact, memory for familiar faces has been demonstrated to be extremely stable over time. For example, one study found that people were able to accurately identify over 90% of high school classmates (irrespective of class size) 15 years after graduation (Bahrick, Bahrick, & Wittlinger, 1975).

Concern regarding the ability to generalize the results from recognition memory paradigms to the eyewitness field has been questioned because identification accuracy rates can change dramatically depending on the experimental design and ecological validity of the tasks (Deffenbacher, Bornstein, Penrod, & McGorty, 2004; Shapiro & Penrod, 1986). Thus, it is essential to investigate the impact of time delays on identification accuracy in an eyewitness context.

In line with the research implicating a power function of memory decay, there is a general consensus among the experts in the eyewitness field that memory drops off quickly over time, followed by a levelling off (Kassin, Ellsworth, & Smith, 1989; Kassin, Tubb, Hosch, & Memon, 2001). In addition, in a review of the eyewitness literature, it was concluded that eyewitness memory loss follows the forgetting curve of Ebbinghaus (Penrod, Loftus, & Winkler, 1982). A meta-analysis on face recognition (which included both recognition and eyewitness paradigms) found that longer delays significantly decreased hit rates, but had less of an impact on false alarms (Shapiro & Penrod, 1986).

In further support of the negative impact of time delay on accuracy, one of the five criteria that arose as a result from the *Neils v. Biggers* (1972) case stated that jurors should take note of the time between the crime and the identification procedure when determining an eyewitness' accuracy. For this case, the accused was charged for rape based solely on a showup identification (i.e., the suspect was presented with no other lineup members present) made seven months after the offence. This criterion implied that accuracy decays as the delay period increases.

In an archival analysis of actual criminal cases, it was found that identifications (corroborated by strong physical trace evidence) were unaffected by delays up to 7-days between the event and the lineup procedure (average identification accuracy rate was 55%), but accuracy rates dropped (average 45%) for delays longer than 7-days (Behrman & Davey, 2001).

A serious concern in the eyewitness field is the confounding effect of interference that may occur during the time delay between witnessing a crime and the identification procedure. There is now a substantial body of literature demonstrating that numerous factors and variables occurring after the event and prior to the lineup (e.g., the number of intervening photographs, leading questions and misinformation) can negatively influence accuracy rates (see Wells & Loftus, 2003; Wells & Olsen, 2003). It appears that these factors contribute to the negative view that memory deteriorates quickly over time in eyewitness contexts. Therefore, there is a need for studies investigating the impact of time delays while controlling for other interfering variables.

Despite the apparent consensus regarding the negative impact of time delay, there has been limited eyewitness research directly comparing identification rates across multiple retention intervals. Currently, there have been only a limited number of experimental studies that have specifically investigated face identification accuracy across more than one time delay within the same study. One early study investigated the impact of multiple (between subjects) time delays (4-minute, 30-minute, 1-hour, 4-hour, 1-day and 1-week) on face recognition (Fessler, Lenorovitz, & Yoblick, 1974). For this study, participants either viewed four different photographs of a target individual or watched a 1-minute film containing the target. A series of 149 sequential photographs (one of which was the target) were shown to participants following one of the six time delays. Of most interest, there was no significant impact of time delay on identification accuracy. They also found that identification accuracy was lower for participants that watched the film compared to the photographs of the target. These results provide further evidence that the type of exposure to the target plays a role in accuracy. This result suggests that accuracy rates may be lower in an eyewitness context compared to face recognition studies that repeat the same photographs in both the encoding and recognition phases.

In another study (Shepherd, Davies, & Ellis, 1980), participants watched a staged incident followed by a live lineup after a delay of 1-week, 1-month, 3-month or 1-year. Identification accuracy was found to be highly stable up until 1-month, but then dropped off to chance levels by the 1-year interval. Furthermore, a study investigated eyewitness accuracy rates in an elderly compared to a young adult population following either 35 minute or 1-week delay. It was found that the elderly participants demonstrated a significant decrease in accuracy at the 1-week delay interval (particularly if the perpetrator was a young adult rather than within their own age category). However, there was no significant decrement in accuracy for the young adult group at the 1-week delay (Memon, Bartlett, Rose, & Gray, 2003).

In two field experiments, no difference in face identification accuracy was found for CP lineups with a 2-minute delay compared to a 4-hour delay (Yarmey, 2004) or from a 2-hour delay compared to a 24-hour delay (Krafka & Penrod, 1985). Similarly, no difference was found for CA lineup accuracy in the study by Yarmey (2004), however, false identifications were found to be substantially increased at the 24-hour delay compared to the 2-hour delay in the study by Krafka and Penrod (1985). In line with these findings, a study with longer time delays (2, 21 and 56 days) found that correct identification rates remained high and relatively stable (e.g., average 97% accuracy for live lineups and 85% for photographic lineups) across the three delay periods. However, the number of false identifications increased over time, although this trend was not statistically significant (Egan, Pittner, & Goldstein, 1977).

Taken together, these results support a position that facial recognition may be fairly resistant to decay, particularly for up to 1-month after the event, when other intervening variables (e.g., biased instructions, repeated questioning, misleading information or use of different viewing conditions) are controlled (Wells & Murray, 1983). In addition, these studies fail, to some degree, to support the common belief that eyewitness memory declines rapidly as a function of time (Kassin, Ellsworth, & Smith, 1989).

Currently there have been no ERP studies conducted using an eyewitness paradigm, however, a few ERP studies have investigated the impact of different retention intervals on old/new effects in recognition memory tasks. As discussed in the general introduction, two old/new effects have been heavily investigated 1) an early frontal old/new effect, which is believed to be associated with familiarity (i.e., a sense of having seen the item before) and 2) a later parietal old/new effect, which is believed to be associated with explicit recollection of old items (see Friedman & Johnson, 2000; Mecklinger, 2000; Rugg, 1995). Studies to date using word stimuli suggest that the parietal old/new effect is sustained following long retention intervals, but that the frontal old/new effect is not, even after brief delays (e.g., a few minutes to 45 minutes) even though behavioural accuracy is still maintained (Rugg, 1990; Rugg & Coles, 1995; Rugg & Nagy, 1989). However, in a paradigm using pictorial stimuli of objects, animals and faces, Curran and

Friedman (2004) demonstrated the maintenance of both the frontal and parietal old/new effects following a 1-day delay.

In one of the few studies that investigated the impact of time delays on ERP components in a facial recognition task (Joyce & Kutas, 2005), it was found that the parietal old/new effect was, not only present, but demonstrated higher amplitude for correctly remembered faces at the longer retention intervals. Specifically, the positive amplitude was higher at 1-day and 1-week intervals compared to no-delay, 30-minute, and 1-hour delays. The authors interpreted the increased amplitude as a reflection of a stronger memory trace for the items retained over the longer intervals.

Despite some disparities, the results from the ERP literature suggest that the parietal old/new effect underlies long-term memory processes and provides a reliable index of face recognition that is maintained overtime. Furthermore, ERP deception detection studies have been able to distinguish P300 patterns associated with guilty knowledge from unrehearsed past events that may have occurred years earlier (Johnson & Rosenfeld, 1992). Based on these results, it was expected that if recognition of the criminal in an eyewitness paradigm is maintained over long time delays, then ERP components should reflect this recognition.

The current experiment investigated the use of ERPs in an eyewitness context to assess the impact of time delays between when participants viewed a crime until when they completed the ERP-lineup across three conditions, no-delay, 1-hour and 1-week delays. For this study, one delay condition was completed before the next one commenced to prevent interference of the conditions and to allow for within subject comparisons across the time delays.

The three main objectives of this experiment were to determine: 1) if ERP patterns were able to differentiate the criminal from the fillers during a correct identification; 2) if identification accuracy decreased as a function of time delay; and 3) if ERP patterns reflected explicit recognition of the criminal even after longer time delays. It was hypothesized that if participants explicitly remember the identity of the criminal, then a LPC would distinguish the criminal from each of the fillers, at each of the three delay periods. In line with ERP recognition studies that demonstrate a decreased parietal old/new effect when there is no explicit recollection, it was hypothesized that the LPC elicited to the criminal would diminish or be absent if participants are unable to remember or distinguish the criminal from the fillers.

Experiment 2a Methods

Participants

Twenty-nine undergraduate students (19 females) with a mean age of 21.7 years ($SD = 2.5$) were recruited from a departmental subject pool and participated in the study for course credit or $5.00 per hour. The data from five participants (three females and two males) were excluded from the analyses[6]. Participants were fluent English speakers with normal hearing, normal or corrected-to-normal vision and no reported history of neurological or psychiatric conditions. Informed consent was obtained at the beginning of the experiment and participants were debriefed at the end of the experiment. The study was approved by the Dalhousie University Research Ethics Board.

---

[6]The data from two participants (S03 and S17) were discarded due to excessive alpha artefacts as a result of drowsiness; S07 fell asleep and subsequently withdraw from the experiment; S16 failed to return for session 2; and EEG technical difficulties resulted in corrupted data for S19.

Stimuli

Crime videos

Four 90 s simulated non-violent crime scenarios were used for the study. The crime scenario was the same for all four videos. However, the identity of the criminal and victim were different in each of the four videos. All four criminals were Caucasian men in their mid-20's. In brief, a female (the victim) walked into an office and answered the phone. She then left the room. When she was gone, a man (the criminal) walked in the room, stole a laptop computer and left. The face of the man was visible for approximately 15 seconds and there were opportunities to see both frontal and profile views. The woman returns to find that her computer is gone.

ERP-lineups (i.e., lineups presented while participants' EEG was recorded)

Four ERP-lineups (A, B, C and D) were used for this study that corresponded to the four crime videos (A, B, C and D). ERP-lineup A consisted of seven digitized photographs. The photographs were from the shoulders up with the person looking forward against an off-white background. The photographs were formatted to fit on a 21-inch computer monitor and the brightness/contrast was normalized across the images. The photographs consisted of five fillers (F1, F2, F3, F4 and F5), the criminal (C) from video A and the victim (V). ERP-lineups B, C and D also consisted of seven digitized photographs (F1, F2, F3, F4, F5, C and V). The fillers were of different individuals for all four ERP-lineups. For each ERP-lineup, the fillers were selected prior to the study based on both shared attributes (e.g., approximate age, race and hair length) as well as based on written descriptions of the criminal from five participants after viewing the crime scenarios (see Appendix A for a detailed description of filler selection). Five photographs of individuals matching these descriptions were selected from a database of over 500 photographs and acted as the fillers.

Once the lineup members were selected, a mock witness procedure (Doob & Kirshenbaum, 1973) was conducted to assess for potential lineup bias for the four ERP-lineups. In brief, ten participants were provided with the descriptions of the criminal and were asked to try and identify the criminal from the lineup without ever actually seeing the crime video. The lineup is considered biased if the criminal is selected above more frequently than the remaining lineup members. The lineups were determined to be non-biased prior to the commencement of the study (see Appendix B for a description and the results of the mock witness procedure).

Additionally, for each of the ERP-lineups (A, B, C and D), a photograph of another individual (classified as the *innocent suspect*) replaced the photograph of the criminal to form the CA lineups (i.e., lineups where the criminal is not present). In the same manner as the fillers, the innocent suspects for each of the ERP-lineups were selected based on shared attributes (e.g., approximate age, race and hair length) as well as based on written descriptions of the criminal from five participants after viewing the crime scenarios. The innocent suspect was judged to be the individual that looked the most like the criminal by five participants (see Appendix D for a description of how the innocent suspect was selected).

Procedures

The experiment involved four conditions. In the order of presentation, the conditions were 1) 1-hour delay 2) no-delay – criminal present (CP) in the lineup 3) 1-week delay and 4) no-delay - criminal absent (CA) from the lineup (CA condition). The two no-delay conditions were counterbalanced. The criminal was always present in the ERP-lineup for the 1-hour and 1-week delay conditions. For all conditions, prior to watching the video participants were instructed to pay attention and informed they would be asked to try to identify the criminal from a lineup at a later time. After each video, participants were told to pretend that the video they watched was an actual crime, for which they were the only eyewitnesses. They were also asked to pretend that they were at a police station and being asked to try to identify the criminal from a series of photographs and warned that the decisions that they make might lead to the conviction of the person they identify as the criminal. These steps were taken to increase the ecological validity of the task as much as possible. Lastly, participants were told that the *criminal might or might not be present in the lineup.*

Participants were seated in a comfortable chair for the duration of the experiment. Participants attended two sessions, scheduled one week apart.

Session 1: 1-hour delay condition and no-delay (CP or CA) condition

Participants watched one of the four crime videos (counterbalanced across participants and conditions). A 1-hour delay was imposed between watching the first crime video and the presentation of the first ERP-lineup. The electrode set-up occurred during this delay period (discussed below in the Electrophysiological recordings procedures section). The photographs from the ERP-lineup were sequentially and pseudorandomly presented (i.e., all seven photographs were randomly presented prior to their repetition) on a 21-inch computer monitor, which was placed 1.5 m in front of participants at eye level (using Presentation© software). Each of the seven photographs was presented 40 times for 1200 ms with an inter-stimulus interval of 1200 ms.

Following the presentation of each photograph, participants were asked to press one of three buttons on a key pad to identify each photograph as either i) the criminal ii) the victim or iii) an innocent individual (i.e., filler). In contrast to Experiment 1, participants were also given the instructions that *the criminal might or might not be present* in the lineup for all conditions. Thus, these instructions provided participants with the option of not selecting any lineup member as the criminal. Participants were encouraged to respond accurately rather than quickly. The photograph of the victim was used as an experimenter-imposed target to maintain attention and responsiveness to the photographs (similar to catch-trials). In the CA condition and when participants are unable to identify the criminal in the CP lineups, no unique button press responses would be required and hence it was considered important to add a task relevant target (i.e., a photograph of the victim) to ensure engagement in the task. The ERP-lineup took approximately 11 minutes to complete.

Immediately following the ERP-lineup, participants where sequentially shown the photographs again and asked to provide certainty ratings for each of the photographs on a scale of 0 to 10 (where 0 indicated that they were certain it was not the criminal and 10 indicated that they were certain the photograph was of the criminal, and a 5 indicated that they were unsure).

When the ERP-lineup from the 1-hour delay condition was completed, participants watched a second crime video followed by a second ERP-lineup and subsequent certainty ratings (following the same instructions, button press responses and procedures described for the first ERP-lineup). No time delay was imposed, so that the ERP-lineup commenced approximately 1-minute following the completion of the video.

The two no-delay conditions (CP and CA) were counterbalanced, such that half of the participants were given the no-delay CP condition during the second ERP-lineup of session 1, while the other half were presented with the no-delay CA condition. Upon completion of the no-delay condition (either CP or CA), the participants were presented with a third crime video, and informed that the corresponding ERP-lineup would be presented during session 2, following a 1-week delay.

Session 2: 1-week delay condition and no-delay (CP or CA) condition

In session 2, the electrode set-up commenced immediately upon the participant's arrival in the laboratory. Following the electrode set-up, participants were presented with the ERP-lineup that corresponded to the third crime video they watched one week previously at the end of session 1 (i.e., the 1-week delay condition). The same instructions, procedures, button press responses and certainty ratings from the ERP-lineups from session 1 were used during the ERP-lineups on session 2. Following the third ERP-lineup from the 1-week delay condition, participants were presented with a fourth crime video and corresponding ERP-lineup with no-delay (CP or CA, i.e., half of the participants were immediately presented with CP ERP-lineup and the other half were presented with a CA ERP-lineup).

Experiment 2a versus 2b

The focus of this experiment (Experiment 2a) will be on the impact of the time delays in CP lineups (no-delay, 1-hour and 1-week delay conditions). The CA condition will be discussed as compared to the no-delay CP condition in the next chapter entitled Experiment 2b. The no-delay CP and CA conditions were paired for comparison purposes because both involved immediate presentation of the ERP-lineup following the video and therefore would not be confounded by the delay conditions. From here on, only details pertaining to the three CP delay conditions (no-delay CP, 1-hour delay and 1-week delay conditions) will be described. Details pertaining to the CA condition and comparisons to the no-delay CP condition will be discussed in Experiment 2b.

Electrophysiological recordings

The same electrophysiological recording procedures from Experiment 1 were used (refer to the electrophysiological recording section from Experiment 1 methods). Trials with EOG voltages greater than +/- 75 μV were discarded from the analyses. Following EOG artifact rejection, a mean of 89.2% of the data (with a range of 75.4 - 98.9 %) was retained for the analyses in the no-delay condition, 86 % (range 66.1 - 98.9 %) for the 1-hour delay condition and 86.3 % (range 70 – 98.9 %) for the 1-week delay condition.

Grand average waveforms

The ERPs from the EOG artifact-free data for each individual participant were averaged according to photograph (F1, F2, F3, F4, F5 and C) for the no-delay, 1-hour and 1-week delay conditions. In other words, the 40 recordings of each photograph (minus the discarded trials with EOG artifacts) were averaged to create six waveforms for each individual for each condition. Grand (or group) average waveforms were created for each photograph (F1, F2, F3, F4, F5 and C) for each condition by averaging the individual averages.

Experiment 2a Results

Grand average waveforms and ANOVA analyses

Figures 8, 9 and 10 depict the grand average waveforms for the no-delay, 1-hour and 1-week delay conditions, respectively, for the 30 scalp electrodes elicited to the criminal and the five fillers averaged together. The LPC elicited to the criminal compared to the fillers was visually apparent at the centro-parietal electrode sites for each of the three delay conditions. Figures 11, 12 and 13 depict the grand average waveforms for the three conditions, respectively, for the Cp3, Cpz, Cp4, P3, Pz and P4 electrode sites for each of the lineup members. Based on these figures, a distinguishable LPC was elicited to the criminal compared to each of the fillers at all six of the electrode sites between 400-600 ms post-stimulus onset.

Two 3-way repeated measures analysis of variance (RM ANOVA) with CONDITION (no-delay, 1-hour delay and 1-week delay), PHOTO (F1, F2, F3, F4, F5 and C) and SITE (Cp3, Cpz, Cp4, P3, Pz and P4) as factors were conducted. The ANOVA analysis was subjected to a Greenhouse-Geisser conservative degrees of freedom correction (Greenhouse & Geisser, 1959). The first ANOVA utilized the mean amplitude for the 400-600 ms time interval (*fixed-interval analysis*), whereas, the second ANOVA utilized the mean amplitude +/- 50 ms around the peak of the criminal as the dependant variable (*criminal-based analysis*; refer to the Results section of Experiment 1 for more information on these two types of analyses).

For both the fixed-interval and criminal-based analyses, planned comparisons were computed comparing the criminal from the other fillers in each of the three conditions. Only significant main and interaction effects considered germane to the experimental hypotheses were explored (i.e., significant effects collapsed across PHOTO were not explored) and further analyzed by post-hoc comparisons with a Bonferroni correction.

Fixed-interval analyses

The 3-way RM ANOVA (refer to Table 12) indicated a significant main effect of PHOTO. Planned comparisons on this main effect demonstrated that there was increased LPC amplitude elicited to the criminal compared to each of the fillers. There were no significant main or interaction effects with CONDITION, indicating that there were no LPC differences across the three time delay conditions.

The SITE x PHOTO interaction effect revealed that at all six centro-parietal electrode sites, the LPC elicited to the criminal was significantly larger compared to each of the fillers. When comparing across site, the LPC elicited to the criminal demonstrated a midline-right centro-parietal scalp voltage distribution (Pz, P4, Cpz) > (P3, Cp4, Cp3). Planned comparisons also revealed that the mean LPC amplitude elicited to the criminal was significantly greater compared to each of the five fillers at all six electrode sites for each of the delay conditions.

Criminal-based analyses

The RM ANOVA results from the criminal-based analysis (Table 12) yielded almost exactly the same results as the fixed-interval analysis and therefore will not be discussed.

Individual ERP analyses

The same individual ERP analyses described for Experiment 1 was conducted on the three delay conditions for both the fixed-interval and criminal-based analyses.

Additionally, the same criteria for correct identifications, misidentifications and false rejections from Experiment 1 were used (refer to the Individual ERP Analyses section in the Experiment 1 Results).

The individual results for ERP-lineups A, B, C & D for the fixed-interval and criminal-based analyses are presented in Tables 13 and 14 for the no-delay condition, Tables 15 and 16 for the 1-hour delay condition and Tables 17 and 18 for the 1-week delay condition. Figures 14, 15 and 16 depict the individual waveforms at the Pz electrode site for the no-delay, 1-hour and 1-week delay conditions, respectively, for the criminal and fillers averaged together.

Based on visual trends found in Experiment 1, it was speculated that there may be increased LPC amplitude for criminals with distinctive attributes that make them easier to identify relative to faces with more typical characteristics. Participants reports revealed that Criminals A and C were easier to identify compared to Criminals B and D. Therefore, the identification rates from each of the four ERP-lineups will be discussed (although the sample size was too small to conduct statistical analyses).

No-delay condition

i. Fixed-interval analysis

For the no-delay condition, the mean z-score was highest and ≥ 2 to the criminal in 75% (18/24) of participants. Of the remaining six participants (S01, S04, S07, S12, S23 and S29), the mean z-score was highest to one of the fillers, and surpassed the statistical cut off mean z-score of ≥ 2 (i.e., 2 *SD* away from the mean) in four of these participants. Overall, there were a total of 75% (18/24) correct identifications, 17% (4/24) misidentifications and 8% (2/24) false rejections.

When looking at the four ERP-lineups separately, the mean z-score was the highest to the criminal and ≥ 2 compared to the fillers in 4/6 participants for ERP-lineups A and B and in 5/6 participants for ERP-lineups C and D.

ii. Criminal-based analysis

The mean z-score was highest to the criminal in 88% (21/24) of participants and 79% (19/24) had a mean z-score ≥ 2. The remaining three participants (S01, S12, and S23) demonstrated the highest mean z-score to one of the fillers, two of which (S01 and S23) had a mean z-score ≥ 2. Overall, there were a total of 79% (19/24) correct identifications, 8% (2/24) misidentifications and 13% (3/24) false rejections.

When looking at the four ERP-lineups separately, the mean z-score was the highest to the criminal and ≥ 2 compared to the fillers in 5/6 participants for ERP-lineup A and D, 3/6 for ERP-lineup B, and 6/6 for ERP-lineup C.

1-hour condition

i. Fixed-interval analysis

The mean z-score was highest and ≥ 2 to the criminal in 83% (20/24) of participants. The four remaining participants (S06, S12, S18 and S23) demonstrated the highest mean z-score to one of the fillers (for S06 and S23 the mean z-score was ≥ 2). Overall there were 83% (20/24) correct identifications, 8% (2/24) misidentifications and 8% (2/24) false rejections.

When looking at the four ERP-lineups separately, the mean z-score of the criminal was the highest and ≥ 2 compared to the fillers in 6/6 participants for ERP-lineup A, 4/6 participants for ERP-lineup B and 5/6 for ERP-lineups C and D.

ii. Criminal-based analysis

The mean z-score was highest to the criminal in 88% (21/24) of participants (83% or 20/24 had a mean z-score ≥ 2). The remaining three participants (S06, S12 and S23) demonstrated the highest mean z-score to one of the fillers, and of these one (S06) had the mean z-score ≥ 2. Therefore, for the 1-hour condition, there were a total of 83% (20/24) correct identifications, 4% (1/24) misidentifications and 13% (3/24) false rejections.

When looking at the four ERP-lineups separately, the mean z-score was the highest and ≥ 2 to the criminal compared to the fillers in 6/6 participants for ERP-lineup A, 5/6 for ERP-lineup B and D and 4/6 for ERP-lineup C.

1-week condition

i. Fixed-interval analysis

The mean z-score was highest and ≥ 2 to the criminal in 54% (13/24) of participants. The 11 remaining participants (S04, S06, S10, S11, S12, S18, S21, S22, S23, S28 and S29) demonstrated the highest mean z-score to one of the fillers, (7/11 had a mean z-score ≥ 2). Overall, the rate of correct identifications was 54% (13/24), misidentifications 29% (7/24) and 17% (4/24) false rejections.

When looking at the four ERP-lineups separately, the mean z-score was the highest to the criminal and ≥ 2 compared to the other fillers in 4/6 participants for ERP-lineup A, 1/6 for ERP-lineup B, 6/6 for ERP-lineup C and 2/6 for ERP-lineup D.

ii. Criminal-based analysis

The mean z-score was highest to the criminal and ≥ 2 in 58% (14/24) of participants. The remaining ten participants (S04, S06, S10, S11, S12, S21, S22, S23, S28 and S29) demonstrated the highest mean z-score and ≥ 2 to one of the fillers, with the exception of S06, S10 and S28. Overall, the number of correct identifications was 58% (14/24), misidentifications were 29% (7/24) and false rejections were 13% (3/24).

When looking at the four ERP-lineups separately, the mean z-score was the highest to the criminal and ≥ 2 compared to the fillers in 4/6 participants for ERP-lineup A, 1/6 for ERP-lineup B, 6/6 for ERP-lineup C and 3/6 for ERP-lineup D.

Impact of time delay on ERP patterns

Figure 17 depicts the waveforms elicited to the criminal across the three delay conditions in participants that made correct identifications. Although the sample size was too small to conduct statistical analyses, the LPC visually increased for correct identifications made at the longer time delays.

Certainty ratings

The same criteria for correct identifications, misidentifications and false rejections from Experiment 1 were used (refer to Experiment 1 Results, Certainty ratings section). Tables 19, 20 and 21 depict participants' certainty ratings for each of the lineup members for the no-delay, 1-hour and 1-week delay conditions, respectively. Correct identifications based on certainty ratings were 75% (18/24), 92% (22/24) and 58% (14/24), misidentifications were 8% (2/24), 4% (1/24) and 33% (8/24) and false rejections were 17% (4/24), 4% (1/24) and 8% (2/24) for the no-delay, 1-hour and 1-week conditions, respectively. The results demonstrate that correct identifications decreased and misidentifications increased as the delay period increased from 1-hour to 1-week. Interestingly, the correct identification rate was higher in the 1-hour delay compared to the no-delay condition.

Button press accuracy

The same criteria for correct identifications, misidentifications and false rejections from Experiment 1 were used (refer to Experiment 1 Results, Button press response section). Tables 22, 23 and 24 depict the percentage of times participants classified each lineup member as the criminal in the no-delay, 1-hour and 1-week delay conditions, respectively. Correct identifications were 50% (12/24), 83% (20/24) and 50% (12/24), misidentifications were 4% (1/24), 4% (1/24) and 21% (5/24) and false rejections were 46% (11/24), 13% (3/24) and 29% (7/24) for the no-delay, 1-hour and 1-week conditions, respectively. Surprisingly, some participants were inconsistent with their button press responses. Some participants classified more than one lineup member as the criminal (refer to S02 for the 1-hour delay condition and S11, S12, S18, S22 and S29 for the 1-week condition). Other participants only classified a lineup member as a criminal on a limited number of trials (refer to S04, S06, S10, and S22 for the no-delay condition, S18 for the 1-hour delay condition, and S05 for the 1-week delay condition). Lastly, some participants purposely failed to make button responses following each photograph (refer to S04, S07, S20, S25 in Table 22 for the no delay condition, S02, S04, S12, S18 in Table 23, for the 1-hour condition and S11, S12, S25, S29 in Table 24, for the 1-week delay condition). This inconsistent button pressing appears to reflect participant's degree of uncertainty about the identity of the criminal.

Comparison of identification accuracy measures

Tables 25, 26 and 27 provide a summary chart for the number and percentage of correct identifications, misidentifications and false rejections based on certainty ratings, button press responses and ERP mean z-scores for the criminal-based analysis, for the no-delay, 1-hour and 1-week delay conditions, respectively. Generally, the rates of correct identifications based on the three accuracy measures were fairly comparable within each condition. However, button press accuracy was considerably lower compared to the certainty ratings and ERP patterns for the no-delay condition.

A series of Pearson correlations demonstrated statistically strong positive correlations between all three accuracy measures for each of the three delay conditions. Specifically, the percentage of time participants correctly classified the criminal according to button press responses demonstrated strong positive correlations with certainty ratings to the criminal ($r = 0.81$, $r = 0.86$, and $r = 0.90$, $p < 0.001$), for the no-delay, 1-hour and 1-week delay conditions, respectively) and with the mean z-scores to the criminal ($r = 0.65$, $r = 0.67$and $r = 0.82$, $p > 0.001$) for the no-delay, 1-hour and 1-week delay conditions, respectively. The certainty ratings to the criminal also demonstrated a positive correlation with the ERP mean z-scores to the criminal [$r = 0.46$, ($p < 0.05$), $r = 0.60$ ($p < 0.01$) and $r = 0.78$ ($p < 0.001$)] for the no-delay, 1-hour and 1-week delay conditions, respectively. Lastly, for each participant, the percentage of time participants correctly classified the criminal according to button press responses was multiplied by their certainty rating to the criminal. The combined button press and certainty ratings was also found to have strong positive correlations with the ERP mean z-scores to the criminal [$r = 0.57$ ($p < 0.01$), $r = 0.66$ ($p < 0.001$) and $r = 0.81$ ($p < 0.001$)] for the no-delay, 1-hour and 1-week delay conditions, respectively). These results demonstrate that ERP patterns (as indicated by the mean z-scores) are able to accurately reflect identification decisions. Therefore, the higher a participant's percentage of correct button press responses to the criminal and/or the higher their certainty rating to the criminal, then the stronger the ERP differentiation of the criminal from the other lineup members. Conversely, the lower a participant's percentage of correct button press responses to the criminal and/or the lower their certainty rating to the criminal resulted in weaker ERP differentiation of the criminal from the other lineup members.

ERP patterns associated with identification decisions

Figure 18 depicts the waveforms to the criminal collapsed across time delay, for participants that made correct identifications, misidentifications or false rejections based on the individual ERP mean z-score results. Although the sample size was too small to conduct statistical analyses, there was a clear visual trend that demonstrated decreased LPC amplitude to the criminal when participants made false rejections and misidentifications compared to correct identifications. This suggests that identification decisions based on ERP patterns can be visually differentiated from each other.

When examining participants that selected the correct criminal according to both their certainty ratings and button press accuracy, it was found that 95% (all exceeding the mean z-score cut-off of ≥ 2) demonstrated LPC differentiation between the criminal and each of the fillers. More specifically, 92% (11/12), 95% (19/20) and 100% (12/12) of participants had correct identifications based on LPC patterns, for each of the three delay conditions respectively. Therefore, ERPs are able to provide a reliable neurophysiological index of correct identifications when there was a high degree of certainty in participants' identification decision. In addition, when examining participants that made false rejections based on both certainty ratings and button press responses in the no-delay condition, ERP patterns were still able to differentiate the criminal from each of the fillers in 75% (3/4) of these cases.

Experiment 2a Discussion

The purpose of the present study was to investigate the use of ERPs in an eyewitness lineup identification task following three different time delay periods (no-delay, 1-hour delay and 1-week delay) between when participants witnessed the crime video until when they viewed the ERP-lineup. In line with the three main hypotheses of this experiment, it was found that 1) an LPC differentiated the criminal from the fillers; 2) identification accuracy was decreased and was lowest at the longest time delay (i.e., 1-week delay) on an individual level; and 3) LPC amplitude indexed explicit recognition of the criminal, and diminished if participants were unable to correctly identify the criminal.

The major finding was that a centro-parietally based LPC (maximal between 400 and 600 ms post-photograph onset) was elicited to the presentation of the criminal compared to each of the fillers at all three delay conditions. This pattern was statistically very strong on a group level, as demonstrated by the significant differentiation of the criminal from each of the fillers and at each of the six centro-parietal scalp electrodes for all three time delay conditions.

The general consensus amongst eyewitness experts is that eyewitness recognition accuracy declines quickly followed by a gradual leveling off (Kassin, Ellsworth, & Smith, 1989). However, studies directly assessing the impact of delay in eyewitness identification tasks have yielded inconsistent results (Behrman & Davey, 2001; Ebbesen & Rienick, 1998; Egan, Pittner, & Goldstein, 1977; Fessler, Lenorovitz, & Yoblick, 1974; Krafka & Penrod, 1985; Memon, Bartlett, Rose, & Gray, 2003; Penrod, Loftus, & Winkler, 1982; Shapiro & Penrod, 1986; Shepherd, Davies, & Ellis, 1980; Wells & Murray, 1983), with some studies indicating no detrimental impact on accuracy if delays are less than 1-week (Behrman & Davey, 2001) or even up to 1-month (Wells & Murray, 1983).

If the rate of forgetting was greater as a function of increasing time delay, then one might anticipate a diminishing of the LPC on a group level as the retention interval (i.e., time delay) increased. Based on the group analyses, no differences were found across the three conditions. These results indicate that the ERP differentiation of the criminal from the fillers was unaffected by the increase in time delay. However, in a face recognition task, Joyce and Kutas (2005) found an increase in LPC amplitude for explicitly remembered faces at longer time delays (1-week and 1-day versus no-delay, 30 minute and 1-hour delays). The authors interpreted this result as a reflection of a stronger memory trace for accurate recognition after the longer time delays. Therefore, it may be possible that the effect of forgetting at longer time delays may dampen LPC amplitude, but that explicit remembering over the delay may increase LPC amplitude. Taken together, this may lead to an averaging out of these effects on a group level, which may have accounted for the lack of a group LPC delay effect in our study. Thus, it was important to investigate the ERP patterns for individual participants. Furthermore, an analysis on a case-by-case basis is also essential if these methods are to have application in real world settings.

Two types of analyses (fixed-interval and the criminal-based analyses) were conducted for individual participants. The criminal-based analysis led to similar or higher rates of correct identifications based on LPC effects for each of the three conditions. Therefore, the knowledge of the suspect and/or tailoring the analysis to each individual participant provided a narrower latency window (compared to the fixed-interval latency window) that demonstrated a stronger link between LPC amplitude and identification decisions. For this reason, as well as for simplicity, only the criminal-based analysis results will be discussed.

The individual ERP analyses revealed that the criminal was distinguished from each of the fillers in 88%, 88% and 58% of participants for the no-delay, 1-hour delay and 1-week delay conditions, respectively. When looking only at participants whose mean z-score to the criminal was equal to or surpassed the predetermined statistical cut-off score of 2 (i.e., ≥ 2 *SD* from the mean), the criminal was distinguished from each of the fillers (i.e., correct identifications) in 79%, 83%, and 58% of participants, for each of the three delay conditions, respectively.

From these individual results, it was apparent that correct identifications based on LPC patterns were lower in the 1-week condition compared to the no-delay and 1-hour delay conditions. Results from the accuracy rates based on certainty ratings and button press responses also demonstrated a decrease in accuracy from the no-delay and 1-hour delay conditions compared to the 1-week condition. According to participants' certainty ratings, 75%, 92% and 58% of participants made correct identifications (i.e., rated the criminal with the highest certainty rating compared to the fillers), respectively for the three delay conditions. Misidentifications were 8%, 4% and 33%, and false rejections were 17%, 4% and 8%, for the three delay conditions, respectively. For button press responses, correct identifications were 50%, 83% and 50%, misidentifications were 4%, 4% and 21% and false rejections were 46%, 13% and 29% for the no-delay, 1-hour and 1-week conditions, respectively.

The overall pattern that emerged from the three accuracy measures (ERP mean z-scores, certainty ratings and button press responses) demonstrated a decrease in identification accuracy at the 1-week time delay, particularly compared to the 1-hour delay condition. This result was attributed to increased difficulty in explicitly remembering the identity of the criminal following the 1-week delay. Two past studies found that longer time delays did not affect correct identifications but that false identifications and misidentifications were higher at lengthier delays (Egan, Pittner, & Goldstein, 1977; Krafka & Penrod, 1985). In line with these past studies, higher rates of misidentifications were also demonstrated in the current study in the 1-week condition compared to both the no-delay and 1-hour delay conditions.

The last major question posed by this experiment was to determine if the LPC reflected explicit recognition of the criminal. It was hypothesized that LPC amplitude would be reduced if participants were unable to recognize the criminal or were unsure of their judgment. Impressively, by investigating participants that made correct identifications based on both certainty ratings and button press accuracy, it was found that 95% (all exceeding the mean z-score cut-off of ≥ 2) demonstrated LPC differentiation between the criminal and each of the fillers. More specifically, 92%, 95% and 100% of participants had correct identifications based on LPC patterns, for each of the three delay conditions, respectively. This suggests that when participants demonstrate a high degree of accuracy based on certainty ratings and button press responses, there is a very strong likelihood of LPC differentiation of the criminal. Thus, this evidence provides convincing support that ERP patterns can provide an objective index of correct identifications. Furthermore, when participants were unable to explicitly recognize the criminal or demonstrated doubt about the criminal's identity (e.g., misidentifications or false rejections), there was a visible trend showing decreased LPC amplitude (however, the number of participants with misidentifications and false rejections were too low to conduct statistical procedures).

In further support for the relationship between the LPC and explicit recognition, statistically significant positive correlations (ranging from 0.46 to 0.9) were found among 1) the percentage of time participants correctly classified the criminal according to button press responses 2) the certainty ratings to the criminal and 3) the mean z-scores to the criminal for all three time delay conditions. In addition, the percentage of time participants correctly classified the criminal according to button press responses was multiplied by their certainty rating to the criminal and was also found to have strong positive correlation with the ERP mean z-scores to the criminal. In other words, the higher a participant's percentage of correct button press responses to the criminal and/or the higher their certainty rating to the criminal then the stronger the ERP differentiation of the criminal from the other lineup members. Conversely, the lower a participant's percentage of correct button press responses to the criminal and/or the lower their certainty rating to the criminal resulted in weaker ERP differentiation of the criminal from the other lineup members. These results provide evidence that ERP differentiation of the criminal acts a strong indicator of a correct identification.

Taken together, the major findings from this study provide evidence that ERP patterns are strongly associated with participants' identification accuracy. This effect was seen at the individual subject level and not merely for grouped data. LPC patterns were able to reliably differentiate the criminal from the fillers when participants were able to explicitly recognize the criminal but LPC amplitude diminished when participants were uncertain or unable to make a correct identification (e.g., false rejections or misidentifications). Specifically, when participants were able to accurately recognize the criminal, ERP patterns were able to provide a neurophysiological index reflecting the correct identification in 95% of participants across the three delay conditions. Overall, the results of this study are promising and suggest that future research in this area is worthwhile.

Chapter 4: Experiment 2b: Criminal-Present and Criminal-Absent Conditions

Experiment 2b Introduction

Wrongful convictions resulting from false identifications are a major concern in the eyewitness field. It is clear that CA lineups account for a large proportion of the false identifications in real world cases (Clark & Tunnicliff, 2001; Wells, 1993; Wells & Lindsay, 1980; Wells & Turtle, 1986). As discussed in the general introduction, research has demonstrated that eyewitnesses are heavily reliant on the use of a relative judgment strategy. In other words, participants select the individual who looks most like the criminal relative to the other lineup members. Obviously this is of great concern in CA lineups because one of the lineup members will always resemble the actual criminal more than the other lineup members. In addition, if one is relying on a relative judgment strategy, there is no option for indicating that the criminal is not present. In this regard, every time a suspect is actually innocent, there is a risk of a false identification. The risk is inversely proportional to the number of lineup members (e.g., in a 6-member, 1-suspect CA lineup, there is a 17% (1/6) risk of a false identification).

Although the number of CA lineups in actual criminal cases is unknown, they undoubtedly happen (e.g. Wells & Lindsay, 1980; Wells & Turtle, 1986). As a result of the high rates of false identifications that are thought to arise from CA lineups, it has been suggested that a set of evidence-based criteria should be developed and met before a suspect can be placed in a lineup. As an alternative, some authors suggest presenting eyewitnesses with a blank lineup (i.e., a lineup with no suspects) prior to the actual lineup, to rule out possible response bias. If the participant is able to indicate that the criminal is absent, then they may proceed to the actual lineup (Wells & Turtle, 1986).

Another effort to decrease the reliance on a relative judgment strategy and to provide validation for rejecting a lineup has been the introduction of lineup instructions that explicitly warn eyewitnesses that the criminal might or might not be present in the lineup. Malpass and Devine (1981) were the first to demonstrate the impact of these instructions on accuracy for both CA and CP lineups. For CA lineups, false identifications and misidentification rates combined were 78% for the group not given the warning instructions. However, false identifications and misidentifications dropped to 33% when the explicit warning instructions and an option to indicate that *the criminal is not there* were given. Importantly, for CP lineups, the inclusion of the warning instructions had a negligible impact on correct identification rates, demonstrating that the instructions do not simply make participants less likely to select a criminal.

Since this seminal study, these results have been replicated and a comprehensive meta-analysis (Steblay, 1997) found that the addition of CA warning instructions reduced false identifications and misidentifications by 41.6%, but only decreased correct identification rates by 1.9%. A more recent meta-analysis also found that warning instructions reduced the incidence of false identifications and misidentifications and had only a minor impact on CP correct identifications. However, it was also found that biased instructions (e.g., *select the criminal from the lineup*) increased correct identifications in CP lineups (Clark, 2005). In other words, when people were forced to guess in CP lineups, they were likely to correctly select the criminal.

The main objective of the current study was to determine the ability of ERP patterns to assess identification accuracy in CA compared to CP lineups. The lineup warning instructions were used in the present study. For the CA condition, it was hypothesized that correct rejections (i.e., indicating that the criminal was not in the lineup) would result in indistinguishable LPC patterns across each of the lineup members. This hypothesis was based on prior studies in the ERP recognition memory field that demonstrated a diminished or absent parietal old/new LPC effect to unrecognized items not associated with explicit recollection (e.g. Duzel, Yonelinas, Mangun, Heinze, & Tulving, 1997; Rugg, et al., 1998; Ullsperger, Mecklinger, & Muller, 2000). In addition, in the deception detection ERP studies, the P300 component is not present to the crime-related item in the innocent group, which is interpreted as a lack of knowledge about its significance (e.g. Allen & Iacono, 1997; Farwell & Donchin, 1991; Rosenfeld, 2002; Rosenfeld, Angell, Johnson, & Qian, 1991; Rosenfeld, et al., 1999).

However, the concern that there is always one lineup member that looks the most like the criminal also applies to ERP CA-lineups. From an ERP perspective, there will always be one lineup member that is associated with the highest LPC amplitude (referred to as the *highest CA lineup member*). Thus, there is some concern of false identifications in an ERP-identification task. Therefore, the second main objective of this experiment was to investigate if the LPC effects known to be elicited to correct identifications in CP lineups can be differentiated from the LPC elicited to the highest CA lineup member.

Although there have been no ERP studies using an eyewitness paradigm, there have been a number of recent ERP studies that have attempted to distinguish true from false recognition. A commonly used paradigm to investigate false recognition using ERPs involves an adaptation of the old/new recognition memory paradigms (Deese, 1959; Roediger & McDermott, 1995). Generally, this involves presenting old and new words, as well as a third category of words that are new but are semantically and/or categorically related to the old words (lures). Behavioural studies have found that lures are much more likely to be incorrectly classified as old and hence reflect false recognition (e.g. Roediger & McDermott, 1995).

Using a lure paradigm, ERP studies have been able to distinguish true (correct identification of old items) from false recognition (incorrect classification of lure items as old). It has been argued, that the frontal old/new effect reflects a sense of familiarity associated with an item, whereas the parietal old/new effect (or LPC) is related to recollection (refer to the ERP and recognition memory section of the General Introduction or see Rugg & Yonelinas, 2003). From this perspective, one would expect both old items and lures to demonstrate a frontal old/new component compared to new items, because lures are expected to elicit a sense of familiarity. However, one would expect that recollection of old items would be superior to lures. Hence, it should be feasible to differentiate true from false recognition based on the parietal old/new effect.

Many studies have demonstrated that the frontal old/new effect is indistinguishable between old and lure items compared to new items (Curran, Schacter, Johnson, & Spinks, 2001; Nessler, Mecklinger, & Penney, 2001). However, the results for the parietal old/new effect (or LPC) have been less consistent. Nessler et al. (2001) found increased LPC amplitude for old compared to lure and new words. However, these effects were most prominent in participants that used an encoding strategy based on studying the specific features of the words compared to participants that studied the words based on conceptual similarity. Curran et al. (2001) also found LPC differences between old words compared to lures, but only in poor performers with high rates of false recognition.

Other studies have failed to find LPC differences and instead have been able to distinguish true and false recognition by other ERP components. One study found a decrease in P300 latency for lures compared to old items (Miller, Baratta, Wynveen, & Rosenfeld, 2001). Another study found that an error-related negativity (ERN, a response-locked negative component believed to be associated with response uncertainty) to be elicited to lure items compared to old items (Nessler & Mecklinger, 2003). A larger late frontal negative component (maximal at 750 ms post stimulus onset) was found to be elicited to lures compared to old items (Nessler & Mecklinger, 2003). Lastly, in one of the few studies that looked at false recognition for faces (Endl, Walla, Lindinger, Deecke, & Lang, 1999), a later frontal negative component was also able to distinguish true from false recognition. Therefore, although the exact components are not necessarily replicated across multiple studies, there is enough evidence to expect that different ERP patterns should be able to dissociate explicit recognition of the criminal compared to false identifications.

Thus, in relation to the two main objectives of the present study it is hypothesized that 1) if participants make correct rejections, there will be no distinguishable ERP patterns across any of the lineup members and 2) the highest CA lineup member will be distinguishable from correct identifications in CP lineups. Specifically, it is hypothesized that LPC amplitude elicited to correct identifications in CP lineups will be greater than the amplitude elicited to a false identifications or misidentifications in CA lineups.

Experiment 2b Methods

Participants/stimuli and procedures

The participants, stimuli and procedures for this study have previously been reported in the methods section for experiment 2a. The focus of Experiment 2a was on the impact of the time delays in CP lineups (no-delay, 1-hour and 1-week delay conditions). The focus of the Experiment 2b will be on the CA condition as compared to the no-delay CP condition. The no-delay CP and CA conditions were paired for comparison purposes because both involved immediate presentation of the ERP-lineup following the video and therefore there were no confounding delay effects.

Experiment 2b Results

This experiment focused on the no-delay CA condition in comparison to the no-delay CP condition. The results for the no-delay CP condition have previously been described (refer to the No-delay condition in the Results section for Experiment 2a).

Electrophysiological results

Trials with EOG voltages greater than +/- 75 μV were discarded from the analyses. Following EOG artifact rejection, a mean of 86.1% of the data (range 67.9-98.9%) was retained for the analyses for the CA condition. Grand average waveforms were created for each of the lineup members [F1, F2, F3, F4, F5 and innocent suspect (IS)] in the CA condition.

Figure 19 depicts the grand average waveforms for the CA condition elicited to the innocent suspect and the fillers averaged together at the 30 scalp electrodes in comparison to the grand average waveform to the criminal in the CP condition. Figure 20 depicts the grand average waveforms for each lineup member in the CA condition, for the

Cp3, Cpz, Cp4, P3, Pz and P4 electrode sites as well as the grand average of the criminal in the CP condition for visual comparison. Visual inspection of these waveforms demonstrated that there was no distinguishable LPC pattern differentiating any of the lineup members from each other on a group level.

In line with Experiments 1 and 2a, the fixed-interval and criminal-based analyses were conducted. However, for the CA condition, the criminal was replaced with an innocent suspect. Therefore, in place of the criminal-based analysis, the peak of the innocent suspect (IS) was detected and the mean amplitude +/- 50 ms around this peak served as the reference latency window for each participant. This analysis will therefore be referred to as the *IS-based analysis*.

Fixed-interval analysis

To investigate the LPC effects in the CA condition, a 2-way RM ANOVA was conducted with PHOTO (F1, F2, F3, F4, F5 and IS) and SITE (Cp3, Cpz, Cp4, P3, Pz and P4) as factors. The analysis was subjected to a Greenhouse-Geisser conservative degrees of freedom correction (Greenhouse & Geisser, 1959). Of most interest, there was no significant PHOTO main effect, indicating that none of the lineup members (including the innocent suspect) were significantly differentiated from each other on a group level (refer to Table 28).

To investigate the LPC differences between the CA and CP (i.e., the no-delay condition from Experiment 2a) conditions, a three-way RM ANOVA was conducted with CONDITION (CP and CA), PHOTO (F1, F2, F3, F4, F5 and C/IS) and SITE (Cp3, Cpz, Cp4, P3, Pz and P4) as factors. The analysis was subjected to a Greenhouse-Geisser conservative degrees of freedom correction (Greenhouse & Geisser, 1959). Relevant significant main effects and interaction effects were further analyzed by planned comparisons or post-hoc comparisons with a Bonferroni correction. Only significant main and interaction effects considered germane to the experimental hypotheses were explored (i.e., significant effects collapsed across PHOTO or CONDITION were not explored). The results (refer to Table 29) are best interpreted within the significant CONDITION x SITE x PHOTO interaction. Therefore, post hoc tests were computed only on this interaction effect.

For the CA condition, there was no significant difference among any of the lineup members, including the innocent suspect. In addition, when comparing across the CA and CP conditions, the LPC elicited to the criminal in the CP condition was significantly larger compared to the innocent suspect, or any other filler in the CA condition. These results demonstrated that when the criminal was absent from the lineup, the LPC elicited to the criminal in the CP condition was not elicited in the CA condition.

For the CA condition, there will always be one lineup member that elicits the highest amplitude LPC within the temporal window of interest (referred to as the highest CA lineup member). Thus, it was considered important to determine if the LPC elicited to the highest CA member could be reliably differentiated from the LPC elicited to a correct identification in the CP condition. To assess this, a 2-way RM ANOVA was conducted with CONDITION (CP and CA) and SITE (Cp3, Cpz, Cp4, P3, Pz and P4) as factors. Only the criminal in the CP condition and the highest CA lineup member in the CA condition were included in the analyses, hence there was no PHOTO factor. In addition, only participants that made correct identifications in the CP condition (based on the fixed-interval individual analysis, N=18) were included in the analyses. Post-hoc comparisons with a Bonferroni correction were conducted on relevant significant main and interaction effects. The results (see Table 30) demonstrated that the LPC elicited to the criminal in participants that made correct identifications in the CP condition was significantly greater compared to the LPC elicited to the highest CA lineup member. Post hoc multiple comparisons on the significant CONDITION x SITE effect demonstrated that the LPC elicited to the criminal in the CP condition was significantly greater than to the highest CA member at all six electrode sites. Therefore, the results demonstrate that the LPC associated with correct identifications can be reliably distinguished from the LPC elicited to the highest CA lineup member on a group level.

IS-based analysis

The same three RM ANOVA analyses conducted for the fixed-interval analyses (described above) were conducted for the IS-based analysis (refer to Tables 28, 29 and 30). The group results from the criminal/IS-based analyses were almost identical to the fixed-interval analysis and therefore will not be discussed.

Individual ERP analyses

The same individual ERP analyses, based on both the fixed-interval and criminal-based latency windows described for Experiment 1 were conducted on the data from the CA condition (refer to the Individual ERP Analysis section in the Experiment 1 Methods), with the exception that the photograph of the criminal was replaced with an innocent suspect. The innocent suspect was judged by five participants to be the most similar in appearance to the criminal compared to the other lineup members (refer to Appendix A). Correct identifications, misidentifications and false rejections based on ERP mean z-scores were determined by the same methods described in Experiment 1

Method Identification Accuracy section. However, for the CA condition there was the additional possibility of having correct rejections and false identifications. In the CA condition, a *correct rejection* was assigned if participants' mean z-scores for each of the lineup members (including the innocent suspect) were < 2. A *misidentification* was assigned if the mean z-score to one of the fillers was ≥ 2. Lastly, a *false identification* was assigned if the mean z-score associated with the innocent suspect was ≥ 2.

Figure 21 depicts the waveforms elicited to the fillers, the highest CA member and the criminal from the CP condition for each individual participant that made correct identifications. In the instances where participants made misidentifications or false identification in the CP condition, the grand average waveform to the criminal in the CP condition was added for visual comparison.

Fixed-interval analysis

Table 31 depicts the individual results for the fixed-interval analysis for ERP-lineups A, B, C & D for the CA condition. The results demonstrated that 42% (10/24) of participants made correct rejections. The 14 remaining participants (S02, S04, S06, S07, S09, S11, S13, S14, S20, S21, S23, S24, S28, and S29) had a mean z-score ≥ 2 to one of the other lineup members. More specifically, 46% (11/24) made misidentifications (i.e., the highest CA lineup member was a filler) and 13% (3/24) made false identifications (i.e., the highest CA lineup member was the innocent suspect). When looking at the four ERP-lineups separately, 2/6 participants for ERP-lineups A and C and 3/6 participants for ERP-lineups B and D demonstrated ERP patterns indicative of correct rejections.

A pair-wise t-test was conducted that compared the mean z-scores elicited to the criminal in the CP condition for participants that made correct identifications compared to the mean z-scores elicited to the highest CA lineup member in the CA condition. In line with the group results from the 2-way RM ANOVA, the mean z-scores associated with the criminal in the CP condition (mean = 4.82) were significantly higher compared to the highest CA lineup member (mean = 2.94, $p < 0.001$). In addition, 83% (15/18) of participants had a larger mean z-score to a correct identification of the criminal in the CP condition compared to the highest CA lineup member.

IS-based analysis

Table 32 depicts the individual results for the IS-based analysis for ERP-lineups A, B, C & D for the CA condition. The results demonstrated that 46% (11/24) of participants made correct rejections. The 13 remaining participants (S02, S04, S05, S09, S11, S13, S14, S15, S20, S21, S24, S28, and S29) had a mean z-score ≥ 2 to one of the lineup members. More specifically, 42% (10/24) made misidentifications (i.e., the highest CA lineup member was a filler) and 13% (3/24) made false identifications (i.e., the highest CA lineup member was the innocent suspect). When looking at the four ERP-lineups separately, 2/6 participants for ERP-lineups A and D, 4/6 participants for ERP-lineup B and, 3/6 participants for ERP-lineup C, demonstrated ERP patterns indicative of correct rejections.

A pair-wise t-test was conducted that compared the mean z-scores elicited to the criminal in the CP condition for participants that made correct identifications compared to the mean z-scores elicited to the highest CA lineup member in the CA condition. Once again, the mean z-scores associated with the criminal in the CP condition (mean = 4.97) were significantly higher compared to the highest CA lineup member (mean = 2.72, $p<0.001$). In addition, 79% (15/19) of participants had a larger mean z-score to a correct identification of the criminal in the CP condition compared to the highest CA lineup member.

Identification accuracy

Correct identifications, misidentifications and false rejections based on certainty ratings and button press responses were determined by the same methods described in Experiment 1 Method Identification Accuracy section. However, for the CA condition there was the additional possibility of having a correct rejection and a false identification. For certainty ratings, a *correct rejection* was classified if participants failed to rate any of the lineup members with a certainty rating of ≥ 5. For button press responses, a *correct rejection* was defined if no lineup member was classified as the criminal more than 50% of the time. For certainty ratings, a *false identification* was classified if participants ranked the photograph of the innocent suspect as the highest certainty rating and ≥ 5. For button press responses, a *false identification* was classified if participants classified the innocent suspect as the criminal most frequently and at least 50% of the time.

Certainty ratings

Table 33 depicts participants' certainty ratings for each of the lineup members in the CA condition. Based on certainty ratings, 50% (12/24) of participants made correct rejections and 50% (12/24) made misidentifications. No participants made false identifications.

Button press accuracy

Table 34 depicts the percentage of times participants classified each lineup member as the criminal in the CA condition. Based on button press accuracy, 83% (20/24) of participants made correct rejections, 17% (4/24) made misidentifications and

0% (0/24) made false identifications. Five participants (S04, S06, S12, S21 and S28) demonstrated some inconsistency with their button press responses, suggesting some degree of uncertainty about their responses.

Comparison of identification accuracy measures

Table 35 is a summary chart that compares the number and percentages of correct rejections, misidentifications and false identifications across the three accuracy measures (certainty ratings, button press responses and ERP patterns from the innocent suspect-based analysis), respectively. Interestingly, the number of correct rejections was higher based on the button press responses (83%) compared to the certainty ratings (50%) and ERP patterns (46%).

According to all three accuracy measures (ERP mean z-scores, certainty ratings and button press responses), participants that made identification decision errors (i.e., a false identification or a misidentification), were not more likely to incorrectly select the innocent suspect as the criminal compared any of the other lineup members. This finding most likely reflects the high degree of similarity and overlapping features of each lineup member with the actual criminal, such that one of the lineup members (including the innocent suspect) does not stand out as more closely resembling the actual criminal across participants.

Experiment 2b Discussion

The main objectives of this experiment was to investigate ERP patterns when the criminal was absent from the lineup. It was hypothesized that the LPC previously demonstrated to differentiate the criminal from the other lineup members would not be present in the CA condition. Therefore, it was expected that there would be no distinguishable LPC elicited to any of the lineup members, including the innocent suspect. In agreement with the hypothesis, the group analyses demonstrated that none of the lineup members in the CA condition were significantly differentiated from each other both collapsed across and each of the six centro-parietal sites individually. In addition, the LPC elicited to the criminal in the CP condition was significantly larger compared to each of the lineup members in the CA condition. These group results provide evidence that there was no distinguishable LPC when the criminal was not in the lineup.

However, for application in the eyewitness field, it is important to demonstrate that there is no distinguishable LPC to one of the lineup members in a CA lineup on a case-by-case basis. The individual ERP results for the CA condition (based on the innocent suspect analysis) indicated that 46% of participants made correct rejections. However, in the remainder of the participants, 42% made misidentifications and 13% made false identifications. Accuracy rates based on participants' certainty ratings indicated that 50% made correct rejections and the remaining 50% made misidentifications. According to button press responses, 83% made correct rejections, 17% made misidentifications and 0% made false identifications.

Interestingly, the percentage of correct rejections was higher for button press responses (83%) compared to ERP patterns (46%) and certainty ratings (50%). This is interpreted as a reflection of increased cautiousness for the button press response task. From examining participants' button press responses, a number of participants waited for each lineup member to be presented multiple times before selecting or rejecting a criminal. Notably, the results in Experiment 2a also suggested the use of a conservative button press approach. In Experiment 2a, the conservative approach decreased correct identifications in CP lineups, however, it helped increase correct rejections in the CA lineups.

Correct rejection accuracy rates for CA lineups have been found to vary considerably depending on other experimental variables and study designs. Correct rejection rates generally range from 20-70% when lineup instructions warn that the criminal might or might be present in the lineup (e.g. Memon, Hope, & Bull, 2003; Weber, Brewer, Wells, Semmler, & Keast, 2004; Wells & Hryciw, 1984; Wright & Stroud, 2002). In one meta-analysis of 18 studies, correct rejection rates for CA lineups were 34%, although the use of lineup warnings was not controlled (Deffenbacher, Bornstein, Penrod, & McGorty, 2004). In another meta-analysis, correct rejections were 40% when warning instructions were used but only 22% when the warning instructions were omitted (Clark, 2005).

The main rationale for the use of the ERP technique in an eyewitness context is to provide an objective measure that maps onto identification decisions. The individual results suggest that, although there are unique ERP patterns associated with correct rejections, false identifications and/or misidentifications, there is still a risk of an innocent suspect being falsely accused based on ERP patterns. More specifically, if one is relying on ERP patterns for eyewitness accuracy (based on the individual mean z-score results) then the lineup member that evokes the largest LPC in a CA lineup (i.e., the highest CA lineup member) may be at risk for being misclassified as the criminal.

Based on previous ERP recognition memory studies, the parietal old/new effect is believed to underlie explicit recollection and is larger to correctly remembered items compared to forgotten items (Curran, 1999; Duzel, Yonelinas, Mangun, Heinze, &

Tulving, 1997; Rugg, et al., 1998; Rugg, Schloerscheidt, & Mark, 1998). In addition, in ERP false recognition memory paradigms, items incorrectly classified as old can be differentiated from actual old items (Curran, Schacter, Johnson, & Spinks, 2001; Miller, Baratta, Wynveen, & Rosenfeld, 2001; Nessler, Mecklinger, & Penney, 2001). Therefore, ERPs should ideally be able to provide a neurophysiological correlate of explicit recognition of the criminal that should be lacking for the highest CA lineup member.

Thus, the second main objective was to determine if the highest CA lineup member could be differentiated from a correct identification in the CP lineup. To investigate this, participants that made correct identifications in the CP condition had their waveforms to the criminal compared to the highest CA lineup member in the CA condition. It was found that the LPC to correct identifications was significantly larger compared to the highest CA lineup member both collapsed and across all six centro-parietal electrode sites. Moreover, the average mean z-score elicited by correct identifications in the CP condition was significantly greater compared to the highest CA lineup member. The mean z-scores to correct identifications were substantially higher compared to the highest CA lineup member in 83% of participants based on the fixed-interval latency window and 79% based on the innocent-suspect-based analysis. Taken together, these results provide strong evidence that an accurately identified criminal can be reliably distinguished from the highest CA lineup member, both on a group and individual level.

It is believed that the overlap of the LPC elicited to the criminal during a correct identification and the LPC elicited to the highest CA lineup member is due to a targetness and task relevance effect that evolves as the ERP-lineup task progresses. In other words, the lineup member speculated or debated by participants to be the potential criminal, most likely receives additional processing compared to the other lineup members. This evokes a target-like situation that has extensively been demonstrated to elicit a classic P300 response (e.g. Donchin, 1981; Johnson, 1986). Therefore, a P300 response is believed to account, at least partially, for the LPC response elicited to both the criminal in CP lineups and to the highest CA lineup member in CA lineups. However, the larger LPC to correct identifications compared to the highest CA lineup member is attributed to an increased parietal old/new effect reflecting explicit recognition of the criminal.

Even though the highest CA lineup member was significantly differentiated from a correct identification, there was still a high rate of false identifications or misidentifications based on the mean z-score ≥ 2 cut-off. This calls into question the cut-off criteria used in this study and the preceding studies to determine accurate decisions based on ERP patterns. Based on the fixed-interval analysis for the CP condition, the average mean amplitude elicited to a correct identification was 15 μV (collapsed across sites) and the average mean z-score was 4.82, compared to an average mean amplitude of 10.5 μV (collapsed across sites) and an average mean z-score of 2.94 elicited to the highest CA member. Similarly, for the innocent-suspect based analyses, the average mean amplitude elicited to a correct identification was 16 μV (collapsed across sites) and the average mean z-score was 4.97. In contrast, the average mean amplitude to the highest CA lineup member was 10 μV (collapsed across sites) and the average mean z-score was 2.72. Therefore, perhaps a more stringent amplitude or z-score cut-off needs to be established. More research will be needed to help specify an optimal cut-off criterion that is able to produce high rates of correct identifications while minimizing the risk of false identifications.

In real-world cases, it is important to make a distinction between a false identification and a misidentification because a false identification may lead to a wrongful conviction, whereas a misidentification is a known error. In the real world, an innocent suspect is often alleged to have committed the crime because of their resemblance to the criminal. Therefore, the innocent suspect would be at an increased risk of a being falsely identified. Interestingly, participants that made misidentifications or false identifications in the CA lineups were not more likely to incorrectly select the innocent suspect as the criminal compared to any of the other lineup members. This finding most likely reflects the high degree of similarity and overlapping features of each lineup member with the actual criminal. Hence this demonstrates the importance of non-biased lineup member selection. In addition, the actual criminal is consistently correctly selected among all the lineup members in the CP lineups, but for the CA lineups, the selection of the criminal is random. This also highlights that participants are most likely selecting the actual criminal in CP lineups based on recognition memory.

Taken together, the major findings from this study provide evidence that ERP patterns reliably reflect participants' identification accuracy and that the LPC can differentiate patterns associated with correct identifications compared to false identifications or misidentifications in the CA condition.

Chapter 5: Discussion

The major purpose of this thesis was to investigate the use of ERPs in an eyewitness context and to determine the impact of several variables (known to influence eyewitness accuracy) on ERP patterning. In all Experiments involving CP lineups, a centro-parietally based LPC (maximal between 400 and 600 ms post-photograph onset) was elicited to the presentation of the criminal compared to each of the other lineup members. This pattern was replicated and statistically very strong on both group and individual levels for Experiments 1 and 2a.

In addition to this major finding, the group analyses from Experiment 1, demonstrated that an LPC was elicited to the criminal when participants tried to conceal their knowledge of the criminal. In Experiment 2a, the group analyses demonstrated that LPC amplitude to the photograph of the criminal did not change across the three time delay conditions. This indicates no impact of time delay on ERP patterns on the group level. However, on an individual level, correct identifications based on ERP patterns were reduced at the 1-week time delay. Thus, on an individual level there was support for increased forgetting or diminished ability to accurately identify the criminal at the longer time delay. Moreover, in Experiment 2a, the LPC decreased when participants were unable to explicitly remember the criminal's identity, irrespective of time delay. Strong and statistically significant correlations were also found between certainty ratings, the percentage of correct button press responses and ERP mean z-scores to the criminal.

Lastly, in Experiment 2b, none of the lineup members were distinguishable by an LPC in the CA condition on a group level. This demonstrated that the LPC associated with a correct identification in the CP condition was absent when the criminal was not in the ERP-lineup. Furthermore, the LPC pattern elicited to correct identifications in the CP condition was larger and deviated significantly more from the fillers compared to the highest CA lineup member. This result is important because it demonstrates that the LPC can differentiate patterns associated with correct identifications in CP lineups compared to possible false identifications or misidentifications in CA lineups.

Taken together, the results from Experiments 1, 2a and 2b provide strong support that the LPC can provide a neurophysiological index of explicit recognition of the criminal in comparison to the other lineup members. This effect remains strong, irrespective of 1) whether the person attempts to deny recognition of the criminal and 2) the duration of the time delay between seeing the crime video and the ERP-lineup task. In addition the results also demonstrate that larger LPC effects can help to distinguish false or misidentifications that arise from CA lineups compared to correct identifications in CP lineups.

It is postulated that the LPC component elicited to the criminal reflects a contribution of both a P300 response and a parietal old/new effect associated with explicit recognition. The ERP-lineup task was designed to some degree after the ERP deception detection GKT designs. Generally, this type of design involves presenting crime-related stimuli embedded within a set of crime-irrelevant stimuli. The ERP deception studies refer to the component that differentiates the crime-related from the irrelevant stimuli as a P300. This interpretation makes sense in terms of the extensive oddball paradigm literature demonstrating the elicitation of a P300 to rarely occurring, task relevance stimuli.

However, for both the ERP-GKT and ERP-lineup task, the task relevance and meaningfulness of the stimuli (e.g., the crime-related item or criminal) are dependent on the explicit recognition by the participants. It is this explicit recognition aspect that leads to the proposed elicitation of the parietal old/new effect. From this perspective, it is postulated that the LPC elicited to the criminal provides a neurophysiological index of recognition memory processes over and above the P300 response. The LPC elicited to the criminal is topographically similar with latency and amplitude features that overlap with both the P300 and the parietal old/new effect (Paller, Bozic, Ranganath, Grabowecky, & Yamada, 1999; Paller & Kutas, 1992; Ranganath & Paller, 2000). In a convincing study (Smith & Guster, 1993), it was found that there was a unique contribution of both a rare targetness effect (P300) as well as a recognition effect associated with previously learned items (parietal old/new effect) that together accounted for the positive deflections within the 300-1000 ms interval following stimulus onset.

This interpretation (i.e., that the LPC elicited to the criminal involves a contribution of P300 and parietal old/new effects) can also be used to interpret the LPC differences demonstrated between correct identifications and the highest CA lineup member. When a filler or innocent suspect is incorrectly identified as the criminal, it is likely to develop a targetness and task relevance effect that is known to elicit a P300 response in comparison to the other lineup members. However, only the criminal (and not the filler or innocent suspect) should elicit the explicit recognition aspect known to arise from previously learned items. Therefore, the parietal old/new effect should be associated with an accurate and explicit identification of the criminal.

However, there are alternative interpretations for the increased LPC to correct identifications compared to the highest CA lineup member. For example, there may be more inconsistency in the single sweeps or increased response conflict associated with the highest CA lineup member compared to a correct identification. Response conflict has previously been demonstrated to evoke a negative deflection approximately 100 ms following a button press decision (Johnson, Barnhardt, & Zhu, 2003; Nieuwenhuis, Yeung, van den Wildenberg, & Ridderinkhof, 2003; Ullsperger & von Cramon, 2001). Button press reaction time often falls within the latency window of the LPC. Therefore, elicitation of a negative component within 400-600 ms associated with response conflict may decrease the LPC.

In terms of identification decisions, it is interesting to note the high degree of identification accuracy based on certainty ratings in Experiment 1 compared to previous behavioural studies that have used similar methodologies. For Experiment 1, correct identifications based on certainty ratings were 100%. According to influential meta-analyses in the eyewitness literature, correct identification rates for CP lineups were 51% (Haber & Haber, 2001), 45% (Steblay, Dysart, Fulero, & Lindsay, 2003) and 35% (Steblay, Dysart, Fulero, & Lindsay, 2001)[7]. As argued in the discussion section of Experiment 1, the increased accuracy was attributed to the ERP-lineup task because the multiple repetitions of the photographs may have increased participants' degree of certainty and confidence in their decision. The correct identifications based on certainty

[7] Only the results from sequential lineups are reported for Steblay, Dysart, Fulero, & Lindsay (2001), whereas the other meta-analyses results are collapsed across simultaneous and sequential lineup presentations.

ratings from Experiment 2a were also higher than the general results from the field (75%, 92% and 58% for the no-delay, 1-hour delay and 1-week delay conditions respectively).

However, the number of correct identifications from Experiment 2a was substantially lower than for Experiment 1. It is speculated that this difference may be a result of the lineup warnings that the criminal might or might not be present in the lineup. Past ERP literature has found that the inclusion of these instructions does decrease identification accuracy, although not substantially (Steblay, 1997). In addition, the use of biased lineup instructions (e.g., *select the criminal from the lineup*) have been known to increase correct identifications due to correct guesses guided by relative judgment strategies (Clark, 2005). Thus, guessing may have also contributed to the high rate of correct identifications based on certainty ratings in Experiment 1.

A second interpretation for the decreased accuracy for Experiment 2a compared to Experiment 1 may be related to the differences in the studies designs. The videos in the two studies were comparable in terms of crime scenario (e.g., both depicted a theft from an office) as well as similarities in the length and time of exposure to the criminal. However, there were differences in the lineup photographs. In Experiment 1, the photographs of the lineup members were taken from the waist-up versus from the shoulders-up in Experiment 2a. The extra body cues from the photographs in Experiment 1 may have increased identification accuracy. In addition, in an effort to decrease the ERP-lineup time, the duration of each photograph was decreased from 1500 ms in Experiment 1 to 1200 ms in Experiments 2a and 2b. However, because each image was shown 40 times, this slight decrease in duration is unlikely to have had a noticeable impact on identification accuracy.

Identification decisions based on certainty ratings and LPC patterns closely mirrored each other, but button press decisions led to a lower rate of correct identifications and a higher rate of false rejections for the no-delay condition in Experiment 2a (50% correct identifications compared to 75% for certainty ratings and 79% for ERP mean z-scores) and increased correct rejections for the CA condition in Experiment 2b (83% correct rejections compared to 50% for certainty ratings and 46% for ERP mean z-scores). Thus, participants were less likely to implicate a criminal and appeared to use a more conservative response strategy during the button press task. This was a surprising effect and it is unknown why it occurred. The most likely interpretation is that the button press task provided participants with multiple opportunities to change their minds. Participants were told to make their decision about each photograph the first time it was presented. However, several participants responded somewhat inconsistently or waited until each photograph was presented multiple times before making a decision. These tactics may have added to feelings of uncertainty about the criminal's identity, accounting for the more conservative button press results.

The eyewitness tasks used in all the Experiments posed a challenging identification task. The exposure time to the criminal was minimal (e.g., approximately 15 seconds) and the lineup members were stringently selected to be non-biased and to show a strong resemblance to the actual criminal. Given this context, one can expect that the data from the current set of Experiments offers a conservative or under-representation of correct identification decisions that can be detected with ERP patterns. For example, it is believed that increased exposure to the criminal would result in easier recognition, which subsequently is speculated to lead to a stronger LPC effect. In addition, if someone was previously familiar with criminal (e.g., if they were a close friend or someone they wanted to protect, which could be assumed in the deception condition), then this is also speculated to result in an increased LPC effect.

Videotaped crime scenarios were used in the present study. A common alternative is to conduct staged crimes to increase the ecological validity. One study investigated accuracy for details following a staged robbery compared to a video of the same staged crime (Ihlebæk, Love, Eilertsen, & Magnussen, 2003). Participants that watched the video provided more details and on average had higher accuracy rates. However, both groups demonstrated similar memory and error patterns. The increased accuracy in the video condition was attributed to the increased opportunity to get an overview of the scene in comparison to the staged event. It was concluded that, although laboratory studies may overestimate accuracy rates compared to real crime cases, both staged and video presentations involved similar memory processes and both are able to provide valuable insights into eyewitness memory (Ihlebæk, Love, Eilertsen, & Magnussen, 2003). Moreover, in a meta-analysis, the accuracy rates associated with the use of slide presentations, videos and staged crimes were compared and no large differences were found between these conditions of observation (Lindsay & Harvie, 1988). Taken together, the differences in accuracy between videos and staged crimes appear to be minor. However, it may be beneficial to conduct future research using staged crimes, rather than videos, in an effort to better simulate real-life witnessing conditions.

For the present set of experiments, participants were informed that they would be watching a crime video and would later be asked to identify the criminal from a lineup. Because each participant acted as their own control across conditions it was likely to become obvious after the first condition that the subsequent conditions would involve an identification task. Therefore, participants were told upfront about the lineup task. This knowledge may have altered the way the participants encoded and processed the event. However, a meta-analysis found that prior knowledge about an upcoming recognition task did not have an overall impact on identification accuracy (Shapiro & Penrod, 1986). In addition, Yarmey et al. (2004) directly tested the impact of witness preparation (i.e., awareness of an upcoming identification task) and found no significant impact on recall of criminal-related descriptions and characteristics. Therefore, it is believed that the prior knowledge of the ERP-lineup task in the current set of experiments most likely did not significantly impact the accuracy results.

One unexpected challenge of Experiments 2a and 2b was the classification of correct identifications, misidentifications and correct rejections based on certainty ratings and button press responses. The lack of consistency, and issues about participants changing their minds were not anticipated. Unfortunately, there was no opportunity for participants to give a definite, overall yes, no, or not present response. As a result, there was a need to establish criteria for the classification of identification decisions for both certainty ratings (i.e., a cut-off rating of 5 or greater) and button press accuracy (i.e., classifying a lineup member as the criminal at least 50% of the time). To compensate for this difficulty, future studies should have participants provide a yes or no response for each photograph, as well as the choice to indicate that the criminal is not present.

The study was unable to investigate the impact of the different criminals in the ERP-lineups because of the small sample size that had the same criminal for each condition (n = 10 for Experiment 1, and n = 6 for Experiments 2a and 2b). Although trends were found indicating higher LPC differentiation for the more distinct criminals, it would be worthwhile to investigate the impact of criminal distinctness on ERP patterns in more depth in future studies.

In addition, low rates of incorrect decisions across all the experiments prevented the statistical analysis of ERP patterns associated with correct, false identifications and misidentifications. One way to assess this comparison in future studies would be to greatly increase sample size or to design a study whereby one condition leads to high accuracy, whereas another condition yields low accuracy with each participant as his or her own control. The latter suggestion could be accomplished by the manipulation of variables known to affect identification accuracy (e.g., long exposure times to the criminal compared to short exposure times).

Wrongful convictions are devastating and unjustifiably restrict the freedoms of innocent individuals. The consequences can involve extensive prison terms or even the death penalty (Connors, Lundregan, Miller, & McEwen, 1996; Scheck, Neufeld, & Dwyer, 2000; Wells, et al., 2000; Wells, et al., 1998). It is highly disturbing that faulty eyewitness testimony is the major factor associated with wrongful convictions (Brandon & Davies, 1973; Connors, Lundregan, Miller, & McEwen, 1996; Huff, Rattner, & Sagarin, 1986; Loftus, 1979; Wells, et al., 1998; Woocher, 1977). Despite the errors known to occur with suspect identifications, eyewitness testimony is deemed important evidence. Disregarding or dismissing this type of evidence would most likely lead to increased difficulty convicting criminals, particularly for serious and dangerous offences, and therefore potentially increase the risk to society (Wells & Loftus, 2002; Wells, et al., 2000; Wells & Olsen, 2003).

Currently, with standard eyewitness procedures, it is very difficult, if not impossible, to objectively determine if an eyewitness' identification of a suspect is accurate or not. This thesis investigates the use of ERPs as a potential method to provide an objective measure of eyewitness identification accuracy. Although more research is needed before an ERP-lineup task should be applied to real-world cases, the results of this study are promising and suggest that future ERP research may lead to a tool that can aid and provide neurophysiological verification for identification decisions.

References

Allan, K., & Rugg, M.D. (1997). An event-related potential study of explicit memory on tests of cued recall and recognition. *Neuropsychologia, 35*(4), 387-397.

Allen, J.J., Iacono, W.G., & Danielson, K.D. (1992). The identification of concealed memories using the event-related potential and implicit behavioral measures: a methodology for prediction in the face of individual differences. *Psychophysiology, 29*(5), 504-22.

Allen, J.J.B., & Iacono, W.G. (1997). A comparison of methods for the analysis of event-related potentials in deception detection. *Psychophysiology, 34*(2), 234-240.

Arnoult, M., D. (1956). Familiarity and recognition of nonsense shapes. *Journal of Experimental Psychology, 51*, 269.

Ashcroft, J., Daniels, D., & Hart, S. (2003). *Eyewitness evidence: a trainer's manual for law enforcement*. US Department of Justice, Office of Justice Programs. National Institute of Justice.

Bahrick, H.P., Bahrick, P.O., & Wittlinger, R.P. (1975). Fifty years of memory for names and faces: A cross-sectional approach. *Journal of Experimental Psychology: General, 104*(1), 54.

Barrett, S.E., & Rugg, M.D. (1989). Event-related potentials and the semantic matching of faces. *Neuropsychologia, 27*(7), 913-22.

Barrett, S.E., Rugg, M.D., & Perrett, D.I. (1988). Event-related potentials and the matching of familiar and unfamiliar faces. *Neuropsychologia, 26*(1), 105.

Behrman, B.W., & Davey, S.L. (2001). Eyewitness identification in actual criminal cases: an archival analysis. *Law and Human Behavior, 25*(5), 475-91.

Ben-Shakhar, G., & Furedy, J.J. (1990). *Theories and applications in the detection of deception: A psychophysiological and international perspective.*: Springer-Verlag Publishing.

Bentin, S., & McCarthy, G. (1994). The effects of immediate stimulus repetition on reaction time and event-related potentials in tasks of different complexity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(1), 130-149.

Borchard, E. (1932). *Convicting the innocent: Errors of criminal justice*. Garden City, N.Y. : Garden City Pub. Co.,.

Bothwell, R.K., Brigham, J.C., & Pigott, M.A. (1987). An exploratory study of personality differences in eyewitness memory. *Journal of Social Behavior & Personality, 2*(3), 335-343.

Brandeis, D., & Lehmann, D. (1987). Event-related potentials of the brain and cognitive processes: Approaches and applications. *Neuropsychologia, 12*(5), 98-101.

Brandon, R., & Davies, C. (1973). *Wrongful imprisonment.* London: Allen and Unwin.

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), 337-349.

Brooks, N. (1983). *Police guidelines: Pretrial eyewitness identification procedures.* Ottawa, ON: Law Reform Commission of Canada.

Chance, J., Goldstein, A.G., & McBride, L. (1975). Differential experience and recognition memory for faces. *Journal of Social Psychology*, *97*(2), 243.

Chiappa (1997). *Evoked potentials in clinical medicine.* (3rd ed.). New York: Lippincott-Raven.

Clark, S., & Tunnicliff, J.L. (2001). Selecting lineup foils in eyewitness identification experiments: Experimental control and real-world simulation. *Law and Human Behavior*, *25*(3), 199-216.

Clark, S.E. (2005). A re-examination of the effects of biased lineup instructions in eyewitness identification. *Law and Human Behavior*, *29*(5), 575-604.

Clark, S.E., & Gronlund, S.D. (1996). Global matching models of recognition memory: How the models match the data. *Psychonomic Bulletin & Review*, *3*(1), 37.

Connolly, J.F., D'Arcy, R.C., Lynn Newman, R., & Kemps, R. (2000). The application of cognitive event-related brain potentials (ERPs) in language-impaired individuals: review and case studies. *Int J Psychophysiol*, *38*(1), 55-70.

Connors, E., Lundregan, T., Miller, N., & McEwen, T. (1996). *Convicted by Juries, Exonerated by Science: Case Studies in the Use of DNA Evidence to Establish Innocence After Trial*: U.S. Department of Justice, Office of Justice Programs, National Institute of Justice.

Cory, P. (2001). *The inquiry regarding Thomas Sophonow: The investigation, prosecution and consideration of entitlement to compensation. Retrieved online at: www.gov.mb.ca/justice/sophonow/recommendations/english.html*. Winnipeg, MB Department of Justice.

Cross, T.P., & Saxe, L. (1992). A critique of the validity of polygraph testing in child sexual abuse cases. *Journal of Child Sexual Abuse*, *1*(4), 19.

Curran, T. (1999). The electrophysiology of incidental and intentional retrieval: ERP old/new effects in lexical decision and recognition memory. *Neuropsychologia*, *37*(7), 771-785.

Curran, T. (2000). Brain potentials of recollection and familiarity. *Memory & Cognition, 28*(6), 923-938.

Curran, T., & Friedman, W.J. (2004). ERP old/new effects at different retention intervals in recency discrimination tasks. *Cognitive Brain Research, 18*(2), 107-120.

Curran, T., Schacter, D.L., Johnson, M.K., & Spinks, R. (2001). Brain potentials reflect behavioral differences in true and false recognition. *J Cogn Neurosci, 13*(2), 201-16.

Curran, T., Tanaka, J.W., & Weiskopf, D.M. (2002). An electrophysiological comparison of visual categorization and recognition memory. *Cognitive, Affective & Behavioral Neuroscience, 2*(1), 1-18.

Cutler, B.L., Berman, G.L., Penrod, S., & Fisher, R.P. (1994). Conceptual, practical, and empirical issues associated with eyewitness identification test media. In D.F. Ross, J.D. Read, & M.P. Toglia (Eds.), *Adult eyewitness testimony: Current trends and developments.* (pp. 163-181): Cambridge University Press.

Cutler, B.L., & Penrod, S.D. (1989a). Forensically relevant moderators of the relation between eyewitness identification accuracy and confidence. *Journal of Applied Psychology, 74*(4), 650-652.

Cutler, B.L., & Penrod, S.D. (1989b). Moderators of the confidence^accuracy correlation in face recognition: The role of information processing and base-rates. *Applied Cognitive Psychology, 3*(2), 95-107.

Cutler, B.L., & Penrod, S.D. (1995). *Mistaken identification: The eyewitness, psychology, and the law.*: Cambridge University Press.

Cutler, B.L., Penrod, S.D., & Dexter, H.R. (1990). Juror sensitivity to eyewitness identification evidence. *Law and Human Behavior, 14*(2), 185-191.

Cutler, B.L., Penrod, S.D., & Stuve, T.E. (1988). Jury decision making in eyewitness identification cases. *Law and Human Behavior, 12*, 41-55.

D'Arcy, R.C., Connolly, J.F., & Eskes, G.A. (2000). Evaluation of reading comprehension with neuropsychological and event-related brain potential (ERP) methods. *J Int Neuropsychol Soc, 6*(5), 556-67.

Davies, G., Ellis, H.D., & Shepherd, J. (1978). Face recognition accuracy as a function of mode of representation. *Journal of Applied Psychology, 63*(2), 180.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 58*, 17.

Deffenbacher, K.A., Bornstein, B.H., Penrod, S.D., & McGorty, E.K. (2004). A meta-analytic review of the effects of high stress on eyewitness memory. *Law Hum Behav*, *28*(6), 687-706.

Deffenbacher, K.A., Carr, T.H., & Leu, J.R. (1981). Memory for words, pictures, and faces: Retroactive interference, forgetting, and reminiscence. *Journal of Experimental Psychology: Human Learning & Memory*, *7*(4), 299-305.

Devlin, L. (1976). *Report to the Secretary of State for the Home Department of the Departmental Committee on Evidence of Identification in Criminal Cases*. London: Her Majesty's Stationary Office.

Donaldson, W. (1996). The role of decision processes in remembering and knowing. *Memory & Cognition*, *24*(4), 523-533.

Donchin, E. (1981). Presidential address, 1980. Surprise!...Surprise? *Psychophysiology*, *18*(5), 493-513.

Doob, A., & Kirshenbaum, H. (1973). Bias in police lineups - partial remembering. *Journal of Police Sciences and Administration*, *1*, 287-293.

Duncan-Johnson, C.C., & Donchin, E. (1979). The time constant in P300 recording. *Psychophysiology*, *16*(1), 53-5.

Duzel, E., Yonelinas, A.P., Mangun, G.R., Heinze, H.J., & Tulving, E. (1997). Event-related brain potential correlates of two states of conscious awareness in memory. *Proceeding of the National Academy of Sciences U S A*, *94*(11), 5973-8.

Ebbesen, E.B., & Rienick, C.B. (1998). Retention interval and eyewitness memory for events and personal identifying attributes. *Journal of Applied Psychology*, *83*(5), 745-762.

Ebbinghaus, H. (1885). *Uber das Gedachtnis* Leipzig Duncker & Humblot.

Egan, D., Pittner, M., & Goldstein, A.G. (1977). Eyewitness Identification: Photographs vs. Live Models. *Law and Human Behavior*, *1*(2), 199-206

Ellwanger, J., Rosenfeld, J.P., Sweet, J.J., & Bhatt, M. (1996). Detecting simulated amnesia for autobiographical and recently learned information using the P300 event-related potential. *Int J Psychophysiol*, *23*(1-2), 9-23.

Ellwanger, J., Tenhula, W.N., Rosenfeld, J.P., & Sweet, J.J. (1999). Identifying simulators of cognitive deficit through combined use of neuropsychological test performance and event-related potentials. *J Clin Exp Neuropsychol*, *21*(6), 866-79.

Endl, W., Walla, P., Lindinger, G., Deecke, L., & Lang, W. (1999). Event-related potential correlates of false recognitions of faces. *Neurosci Lett*, *265*(2), 115-8.

Fabiani, M., Karis, D., & Donchin, E. (1983). P300 and memory: individual differences in the von Restoroff effect. *Psychophysiology*(Abstract 558).

Fanselow, M.S., & Buckout, R.F. (1976). *Nonverbal cueing as a source of biasing information in eyewitness identification testing.* New York: Centre for Responsive Psychology, Brooklyn College C.U.N.Y.

Farwell, L.A., & Donchin, E. (1991). The truth will out: Interrogative polygraphy ("lie detection") with event-related brain potentials. *Psychophysiology, 28*(5), 531-547.

Fessler, P.K., Lenorovitz, D.R., & Yoblick, D.A. (1974). Time delay and similarity effects in facial recognition. *Journal of Applied Psychology, 59*(4), 490.

Fischhoff, B., & Beyth-Marom, R. (1983). Hypothesis evaluation from a Bayesian perspective. *Psychological Review, 90*(3), 239-260.

Frank, J., & Frank, B. (1957). *Not guilty.* Garden City, N.Y.: Doubleday.

Friedman, D., & Johnson, R.J. (2000). Event-related potential (ERP) studies of memory encoding and retrieval: A selective review. *Microscopy Research and Technique, 51*(1), 6-28.

Gillund, G., & Shiffrin, R.M. (1984). A retrieval model for both recognition and recall. *Psychological Review, 91*(1), 1.

Girodo, M., Deck, T.P., & Campbell, K.B. (2002). Event-related potentials reveal the effects of altering personal identity. *Neuroreport, 13*(13), 1595-8.

Going, M., & Read, J.D. (1974). Effects of uniqueness, sex of subject, and sex of photograph on facial recognition. *Perceptual and Motor Skills, 39*(1), 109.

Goldstein, A.G., Chance, J.E., & Schneller, G.R. (1989). Frequency of eyewitness identification in criminal cases: A survey of prosecutors. *Bulletin of the Psychonomic Society, 27*, 71-4.

Greenhouse, S.W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika, 24*, 95-112.

Guillem, F., Bicu, M., & Debruille, J.B. (2001). Dissociating memory processes involved in direct and indirect tests with ERPs to unfamiliar faces. *Brain Res Cogn Brain Res, 11*(1), 113-25.

Haber, R.N. (1970). How we remember what we see. *Scientific American, Vol. 222*(5), 104.

Haber, R.N., & Haber, L. (2001). *A meta-analysis of research on eywitness lineup identification accuracy.* Paper presented at the Annual Convention of the Psychonomics Society, Orlando, Florida.

Harrington, T.J. *v. State of Iowa (2003). The Supreme Court of Iowa. No. 122 / 01-0653. Retrieved online: http://caselaw.lp.findlaw.com/scripts/getcase.pl?court=ia&vol=sc%5C20030226%5C01-0653&invol=1.*

Hastie, R., Landsman, R., & Loftus, E.F. (1978). Eyewitness Testimony: the dangers of guessing. *Jurimetrics Journal, 19,* 1-8.

Hepworth, S.L., Rovet, J.F., & Taylor, M.J. (2001). Neurophysiological correlates of verbal and nonverbal short-term memory in children: repetition of words and faces. *Psychophysiology, 38*(3), 594-600.

Hintzman, D.L., Caulton, D.A., & Levitin, D.J. (1998). Retrieval dynamics in recognition and list discrimination: further evidence of separate processes of familiarity and recall. *Mem Cognit, 26*(3), 449-62.

Huff, R., Rattner, A., & Sagarin, E. (1986). Guilty until proven innocent. *Crime and Delinquency, 32,* 518.

Ihlebæk, C., Love, T., Eilertsen, D.E., & Magnussen, S. (2003). Memory for a staged criminal event witnessed live and on video. *Memory, 11*(3), 319-27.

Jasper, H.H. (1958). The ten-twenty electrode system of the international federation. . *Electroencephalography Clinical Neurophysiology, 22*(2), 182-94.

Johnson, M.M., & Rosenfeld, J.P. (1992). Oddball-evoked P300-based method of deception detection in the laboratory. II: Utilization of non-selective activation of relevant knowledge. *Int J Psychophysiol, 12*(3), 289-306.

Johnson, R., Jr. (1986). A triarchic model of P300 amplitude. *Psychophysiology, 23*(4), 367-84.

Johnson, R., Jr., Barnhardt, J., & Zhu, J. (2004). The contribution of executive processes to deceptive responding. *Neuropsychologia, 42*(7), 878-901.

Johnson, R.J., Barnhardt, J., & Zhu, J. (2003). The deceptive response: Effects of response conflict and strategic monitoring on the late positive component and episodic memory-related brain activity. *Biological Psychology, 64*(3), 217-253.

Joyce, C.A., & Kutas, M. (2005). Event-related potential correlates of long-term memory for briefly presented faces. *J Cogn Neurosci, 17*(5), 757-67.

Karis, D., Fabiani, M., & Donchin, E. (1984). 'P300' and memory: Individual differences in the von Restorff effect. *Cognitive Psychology, 16*(2), 177.

Kassin, S.M., Ellsworth, P.C., & Smith, V.L. (1989). The 'general acceptance' of psychological research on eyewitness testimony: A survey of the experts. *American Psychologist, 44*(8), 1089.

Kassin, S.M., Tubb, V.A., Hosch, H.M., & Memon, A. (2001). On the 'general acceptance' of eyewitness testimony research. *American Psychologist, 56*(5), 405.

Kebbell, M.R. (2000). The law concerning the conduct of lineups in England and Wales: how well does it satisfy the recommendations of the American Psychology-Law Society? *Law Hum Behav, 24*(3), 309-15.

Kleinmuntz, B., & Szucko, J.J. (1982). On the fallibility of lie detection. *Law & Society Review, 17*(1), 85.

Kleinmuntz, B., & Szucko, J.J. (1984). Lie detection in ancient and modern times: A call for contemporary scientific study. *American Psychologist, 39*(7), 766.

Knight, R.G. (1997). *Behavioural Neurology and Neurpsychology.* New York: McGraw Hill.

Krafka, C., & Penrod, S. (1985). Reinstatement of context in a field experiment on eyewitness identification. *Journal of Personality and Social Psychology, 49*(1), 58-69.

Leippe, M.R., & Romanczyk, A. (1989). Reactions to child (versus adult) eyewitnesses: The influence of jurors' preconceptions and witness behavior. *Law and Human Behavior, 13*(2), 103-132.

Light, L.L., Kayra-Stuart, F., & Hollander, S. (1979). Recognition memory for typical and unusual faces. *Journal of Experimental Psychology: Human Learning & Memory, 5*(3), 212.

Lindsay, R.C., Lea, J.A., & Fulford, J.A. (1991). Sequential lineup presentation: Technique matters. *Journal of Applied Psychology, 76*(5), 741-745.

Lindsay, R.C., Lea, J.A., Nosworthy, G.J., & Fulford, J.A. (1991). Biased lineups: Sequential presentation reduces the problem. *Journal of Applied Psychology, 76*(6), 796-802.

Lindsay, R.C., Lim, R., Marando, L., & Cully, D. (1986). Mock-juror evaluations of eyewitness testimony: A test of metamemory hypotheses. *Journal of Applied Social Psychology, 16*(5), 447-459.

Lindsay, R.C., Wells, G.L., & O'Connor, F.J. (1989a). Mock-juror belief of accurate and inaccurate eyewitnesses: A replication and extension. *Law and Human Behavior, 13*(3), 333-339.

Lindsay, R.C.L., & Harvie, V.L. (1988). Hits, false alarms, correct and mistaken identifications: The effect of methods of data collection on facial memory. In M.M. Gruneberg, P.E. Morris, & R.N. Sykes (Eds.), *Practical aspects of memory: Current research and issues, Vol. 1: Memory in everyday life.* (pp. 47-52): John Wiley & Sons.

Lindsay, R.C.L., Wells, G.L., & O'Connor, F.J. (1989b). Mock-juror belief of accurate and inaccurate eyewitnesses: a replication and extension. *Law and Human Behavior*, *13*, 333.

Loftus, E.F. (1979). *Eyewitness testimony.* Cambridge, Mass.: Harvard University Press

Loftus, E.F. (1983). Silence is not golden. *American Psychologist, 38*(5), 564-572.

Loftus, E.F., & Loftus, G.R. (1980). On the permanence of stored information in the human brain. *American Psychologist*, *35*(5), 409-420.

Luus, C.A.E., & Wells, G.L. (1994). The malleability of eyewitness confidence: Co-witness and perseverance effects. *Journal of Applied Psychology*, *79*(5), 714-723.

Lykken, D.T. (1981). *A tremor in the blood.* New York: McGraw-Hill.

Malpass, R.S., & Devine, P.G. (1981). Eyewitness identification: Lineup instructions and the absence of the offender. *Journal of Applied Psychology*, *66*(4), 482-489.

Mecklinger, A. (2000). Interfacing mind and brain: a neurocognitive model of recognition memory. *Psychophysiology*, *37*(5), 565-82.

Memon, A., Bartlett, J., Rose, R., & Gray, C. (2003). The aging eyewitness: effects of age on face, delay, and source-memory ability. *Journal of Gerontololgy, Series B, Psychological Sciences and Social Sciences*, *58*(6), P338-45.

Memon, A., Hope, L., & Bull, R. (2003). Exposure duration: effects on eyewitness accuracy and confidence. *Br J Psychol*, *94*(Pt 3), 339-54.

Miller, A.R. (1999). *P300 amplitude and topography in pseudomemory phenomena. (event related potential, simulated amnesia, false recognition)*, Univ Microfilms International.

Miller, A.R., Baratta, C., Wynveen, C., & Rosenfeld, J.P. (2001). P300 latency, but not amplitude or topography, distinguishes between true and false recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(2), 354.

Miller, A.R., & Rosenfeld, J.P. (2004). Response-Specific Scalp Distributions in Deception Detection and ERP Correlates of Psychopathic Personality Traits. *Journal of Psychophysiology*, *18*(1), 13.

Morgan, C.A., 3rd, Hazlett, G., Doran, A., Garrett, S., Hoyt, G., Thomas, P., Baranoski, M., & Southwick, S.M. (2004). Accuracy of eyewitness memory for persons encountered during exposure to highly intense stress. *Int J Law Psychiatry*, *27*(3), 265-79.

Münsterberg, H. (1908). *On the witness stand.* New York: Doubleday.

Murdock, B.B., Jr. (1965). Signal-detection theory and short-term memory. *J Exp Psychol, 70*(5), 443-7.

National, Research, & Council (2003). The Polygraph and Lie Detection. Committee to Review the Scientific Evaluation on the Polygraph , Division of Behavioral and Social Sciences and Education. National Research Council. The National Academies Press, Available on-line at: http://www.nap.edu/books/0309084369/html. Washington D.C.

Neils vs. Biggers (1972) 409 United States 188.

Nessler, D., & Mecklinger, A. (2003). ERP correlates of true and false recognition after different retention delays: Stimulus- and response-related processes. *Psychophysiology, 40*(1), 146.

Nessler, D., Mecklinger, A., & Penney, T.B. (2001). Event related brain potentials and illusory memories: the effects of differential encoding. *Brain Res Cogn Brain Res, 10*(3), 283-301.

Nessler, D., Mecklinger, A., & Penney, T.B. (2005). Perceptual fluency, semantic familiarity and recognition-related familiarity: an electrophysiological exploration. *Brain Res Cogn Brain Res, 22*(2), 265-88.

Nickerson, R.S. (1968). A note on long-term recognition memory for pictorial material *Psychonomic Science, 11*(2), 58.

Nickerson, R.S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2*(2), 175-220.

Nieuwenhuis, S., Yeung, N., van den Wildenberg, W., & Ridderinkhof, K.R. (2003). Electrophysiological correlates of anterior cingulate function in a go/no-go task: effects of response conflict and trial type frequency. *Cogn Affect Behav Neurosci, 3*(1), 17-26.

Paivio, A. (1969). Mental imagery in associative learning and memory. *Psychological Review, 76*(3), 241.

Paller, K.A., Bozic, V.S., Ranganath, C., Grabowecky, M., & Yamada, S. (1999). Brain waves following remembered faces index conscious recollection. *Cognitive Brain Research, 7*(4), 519-531.

Paller, K.A., Gonsalves, B., Grabowecky, M., Bozic, V.S., & Yamada, S. (2000). Electrophysiological correlates of recollecting faces of known and unknown individuals. *Neuroimage, 11*(2), 98-110.

Paller, K.A., & Kutas, M. (1992). Brain potentials during memory retrieval provide neurophysiological support for the distinction between conscious recollection and priming. *Journal of Cognitive Neuroscience, 4*(4), 375-391.

Park, D.C., Royal, D., Dudley, W., & Morrell, R. (1988). Forgetting of pictures over a long retention interval in young and older adults. *Psychology and Aging*, *3*(1), 94-95.

Parker, J.F., & Carranza, L.E. (1989). Eyewitness testimony of children in target-present and target-absent lineups. *Law and Human Behavior*, *13*(2), 133-149.

Parker, J.F., & Ryan, V. (1993). An attempt to reduce guessing behavior in children's and adults' eyewitness identifications. *Law and Human Behavior*, *17*(1), 11-26.

Parliament, L., & Yarmey, A.D. (2002). Deception in eyewitness identification., *Criminal Justice & Behavior* (Vol. 29, pp. 734-746): Sage Publications.

Penrod, S.D., Loftus, E.F., & Winkler, J. (1982). *The psychology of the courtroom* (Vol. pp. 119-168). New York: Academic Press.

Peters, D.P. (1988). Eyewitness memory and arousal in a natural setting. In M.M. Gruneberg, P.E. Morris, & R.N. Sykes (Eds.), *Practical aspects of memory: Current research and issues, Vol. 1: Memory in everyday life.* (pp. 89-94): John Wiley & Sons.

Piggott, M., & Brigham, J.C. (1985). Relationship between accuracy of prior description and facial recognition. *Journal of Applied Psychology*, *70*, 547-55.

Ranganath, C., & Paller, K.A. (2000). Neural correlates of memory retrieval and evaluation. *Cognitive Brain Research*, *9*(2), 209-222.

Reder, L.M., Nhouyvanisvong, A., Schunn, C.D., Ayers, M.S., Angstadt, P., & Hiraki, K. (2000). A mechanistic account of the mirror effect for word frequency: a computational model of remember-know judgments in a continuous recognition paradigm. *J Exp Psychol Learn Mem Cogn*, *26*(2), 294-320.

Regan (1989). Human Brain Electrophysiology: Evoked Potentials and Evoked Magnetic Fields in Science and Medicine. In Regan (Ed.), *Human Brain Electrophysiology: Evoked Potentials and Evoked Magnetic Fields in Science and Medicine*. New York: Elsevier.

Regina (1986). v. Sophonow (No.2), 1986 CanLII 104 (MB C.A.). Retrieved online: http://www.canlii.ca/mb/cas/mbca/1986/1986mbca10001.html.

Regina (1997). v. McGuinnis, Ballantyne & Ballantyne CanLII 1781 (BC S.C.) 10929 Retrieved from http://www.canlii.ca/bc/cas/bcsc/1997/1997bcsc11363.html.

Reno, J., Fisher, R.C., Robinson, L., Brennan, N., & Travis, J. (1999). *Eyewitness Evidence: a guide for law enforcement*. U.S. Department of Justice, Office of Justice Programs. National Institute of Justice.

Roediger, H.L., & McDermott, K.B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(4), 803.

Rosen, P. (1992). *Wrongful convictions in the criminal justice system* Government of Canada. Retrieved online: http://dsp-psd.communication.gc.ca/Collection-R/LoPBdP/BP/bp285-e.htm.

Rosenfeld, J.P. (2002). Event-related potentials in the detection of deception, malingering, and false memories. In M. Kleiner (Ed.), *Handbook of polygraph testing.* (pp. 265-286). San Diego, CA, US: Academic Press.

Rosenfeld, J.P., Angell, A., Johnson, M., & Qian, J.H. (1991). An ERP-based, control-question lie detector analog: algorithms for discriminating effects within individuals' average waveforms. *Psychophysiology, 28*(3), 319-35.

Rosenfeld, J.P., Ellwanger, J.W., Nolan, K., Wu, S., Bermann, R.G., & Sweet, J. (1999). P300 scalp amplitude distribution as an index of deception in a simulated cognitive deficit model. *International Journal of Psychophysiology, 33*(1), 3-19.

Rosenfeld, J.P., Nasman, V.T., Whalen, R., Cantwell, B., & Mazzeri, L. (1987). Late vertex positivity in event-related potentials as a guilty knowledge indicator: a new method of life detection. *Int J Neurosci, 34*(1-2), 125-9.

Rosenfeld, J.P., Rao, A., Soskins, M., & Reinhart Miller, A. (2003). Scaled P300 scalp distribution correlates of verbal deception in an autobiographical oddball paradigm: Control for task demand. *Journal of Psychophysiology, 17*(1), 14.

Rosenfeld, J.P., Reinhart, A.M., Bhatt, M., Ellwanger, J., Gora, K., Sekera, M., & Sweet, J. (1998). P300 correlates of simulated malingered amnesia in a matching-to-sample task: topographic analyses of deception versus truthtelling responses. *Int J Psychophysiol, 28*(3), 233-47.

Rosenfeld, J.P., Soskins, M., Bosh, G., & Ryan, A. (2004). Simple, effective countermeasures to P300-based tests of detection of concealed information. *Psychophysiology, 41*(2), 205-19.

Rugg, M.D. (1990). Event-related brain potentials dissociate repetition effects of high- and low-frequency words. *Memory & Cognition, 18*(4), 367.

Rugg, M.D. (1995). Event-related potential studies of human memory. In M.S. Gazzaniga (Ed.), *The Cognitive Neurosciences.* (pp. 789-801): The MIT Press.

Rugg, M.D., Allan, K., & Birch, C.S. (2000). Electrophysiological evidence for the modulation of retrieval orientation by depth of study processing. *Journal of Cognitive Neuroscience, 12*(4), 664-678.

Rugg, M.D., & Coles, M.G.H. (1995). ERP studies of memory. *Electrophysiology of mind: Event-related brain potentials and cognition.* (pp. 133): Oxford University Press.

Rugg, M.D., Cox, C.J., Doyle, M.C., & Wells, T. (1995). Event-related potentials and the recollection of low and high frequency words. *Neuropsychologia*, *33*(4), 471-84.

Rugg, M.D., Mark, R.E., Walla, P., Schloerscheidt, A.M., Birch, C.S., & Allan, K. (1998). Dissociation of the neural correlates of implicit and explicit memory. *Nature*, *392*(6676), 595-598.

Rugg, M.D., & Nagy, M.E. (1989). Event-related potentials and recognition memory for words. *Electroencephalography & Clinical Neurophysiology*, *72*(5), 395.

Rugg, M.D., Rugg, M.D., & Coles, M.G.H. (1995). ERP studies of memory. *Electrophysiology of mind: Event-related brain potentials and cognition.* (pp. 133): Oxford University Press.

Rugg, M.D., Schloerscheidt, A.M., & Mark, R.E. (1998). An electrophysiological comparison of two indices of recollection. *Journal of Memory and Language*, *39*(1), 47-69.

Rugg, M.D., & Yonelinas, A.P. (2003). Human recognition memory: A cognitive neuroscience perspective. *Trends in Cognitive Sciences*, *7*(7), 313-319.

Saxe, L. (1991). Science and the CQT polygraph: A theoretical critique. *Integrative Physiological & Behavioral Science*, *26*(3), 223.

Saxe, L. (1994). Detection of deception: Polygraph and integrity tests. *Current Directions in Psychological Science*, *3*(3), 69.

Saxe, L., & Ben-Shakhar, G. (1999). Admissibility of polygraph tests: The application of scientific standards post- Daubert. *Psychology, Public Policy, and Law*, *5*(1), 203.

Scapinello, K.F., & Yarmey, A.D. (1970). The role of familiarity and orientation in immediate and delayed recognition of pictorial stimuli. *Psychonomic Science, Vol. 21*(6), 329.

Scheck, B., Neufeld, P., & Dwyer, J. (2000). *Actual Innocence: five days to execution and other dispatches from the wrongly convicted.* New York: Doubleday.

Schloerscheidt, A.M., & Rugg, M.D. (2004). The impact of change in stimulus format on the electrophysiological indices of recognition. *Neuropsychologia*, *42*(4), 451-466.

Schweinberger, S.R., Pfütze, E.-M., & Sommer, W. (1995). Repetition priming and associative priming of face recognition: Evidence from event-related potentials. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(3), 722-736.

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), 2057-73.

Shapiro, P.N., & Penrod, S. (1986). Meta-analysis of facial identification studies. *Psychological Bulletin, 100*(2), 139-156.

Shepard, R.N. (1967). Recogntion memory for words, sentences and pictures. *Journal of Verbal Learning & Verbal Behavior, 6*(1), 156.

Shepherd, J.W., Davies, G.M., & Ellis, H.D. (1980). *Identification after delay.*

Shepherd, J.W., & Ellis, H.D. (1973). The effect of attractiveness on recognition memory for faces. *American Journal of Psychology, Vol. 86*(3), 627-633.

Smith, M.E., & Guster, K. (1993). Decomposition of recognition memory event-related potentials yields target, repetition, and retrieval effects. *Electroencephalogr Clin Neurophysiol, 86*(5), 335-43.

Southwick, S.M., Morgan, C.A., 3rd, Nicolaou, A.L., & Charney, D.S. (1997). Consistency of memory for combat-related traumatic events in veterans of Operation Desert Storm. *American Journal of Psychiatry, 154*(2), 173-7.

Standing, L. (1973). Learning 10,000 pictures. *Quarterly Journal of Experimental Psychology, Vol. 25*(2), 207.

Steblay, N., Dysart, J., Fulero, S., & Lindsay, R.C. (2001). Eyewitness accuracy rates in sequential and simultaneous lineup presentations: a meta-analytic comparison. *Law and Human Behavior, 25*(5), 459-73.

Steblay, N., Dysart, J., Fulero, S., & Lindsay, R.C. (2003). Eyewitness accuracy rates in police showup and lineup presentations: a meta-analytic comparison. *Law and Human Behavior, 27*(5), 523-40.

Steblay, N.M. (1997). Social influence in eyewitness recall: A meta-analytic review of lineup instruction effects. *Law and Human Behavior, 21*(3), 283.

Tollerstrup, P.A., Turtle, J.W., & Yuille, J.C. (1994). Actual victims and witnesses to robbery and fraud: An archival analysis. In D.F. Ross, J.D. Read, & M.P. Toglia (Eds.), *Adult eyewitness testimony: Current trends and developments.* (pp. 144-160): Cambridge University Press.

Turtle, J.W., Lindsay, D.S., & Wells, G.L. (2003). Best Practice recommendations for eyewitness evidence procedures: new ideas for the oldest way to solve a case. *The Canadian Journal of Police and Security Services.*

Turtle, J.W., & Yuille, J.C. (1994). Lost but not forgotten details: repeated eyewitness recall leads to reminiscence but not hypermnesia. *J Appl Psychol*, *79*(2), 260-71.

Ullsperger, M., Mecklinger, A., & Muller, U. (2000). An electrophysiological test of directed forgetting: the role of retrieval inhibition. *J Cogn Neurosci*, *12*(6), 924-40.

Ullsperger, M., & von Cramon, D.Y. (2001). Subprocesses of performance monitoring: a dissociation of error processing and response competition revealed by event-related fMRI and ERPs. *Neuroimage*, *14*(6), 1387-401.

Wasserman, S., & Bockenholt, U. (1989). Bootstrapping: applications to psychophysiology. *Psychophysiology*, *26*(2), 208-21.

Weber, N., Brewer, N., Wells, G.L., Semmler, C., & Keast, A. (2004). Eyewitness identification accuracy and response latency: the unruly 10-12-second rule. *Journal of Experimental Psychology Applied*, *10*(3), 139-47.

Wells, G.L. (1984a). Do the eyes have it? More on expert eyewitness testimony. *American Psychologist*, *39*(9), 1064-1065.

Wells, G.L. (1984b). The psychology of lineup identifications. *Journal of Applied Social Psychology*, *14*, 89.

Wells, G.L. (1984c). The psychology of lineup identifications. *Journal of Applied Social Psychology*, *14*(2), 89-103.

Wells, G.L. (1985). Verbal descriptions of faces from memory: Are they diagnostic of identification accuracy? *Journal of Applied Psychology*, *70*, 619-26.

Wells, G.L. (1993). What do we know about eyewitness identification? *Am Psychol*, *48*(5), 553-71; discussion 572-80.

Wells, G.L. (2001). Police lineups: Data, theory, and policy. *Psychology, Public Policy, and Law*, *7*(4), 791-801.

Wells, G.L., & Hryciw, B. (1984). Memory for faces: Encoding and retrieval operations. *Memory & Cognition*, *12*(4), 338-344.

Wells, G.L., & Leippe, M.R. (1981). How do triers of fact infer the accuracy of eyewitness identifications? Using memory for peripheral detail can be misleading. *Journal of Applied Psychology*, *66*(6), 682-687.

Wells, G.L., & Lindsay, R.C. (1980). On estimating the diagnosticity of eyewitness nonidentifications. *Psychological Bulletin, 88*(3), 776-784.

Wells, G.L., Lindsay, R.C., & Ferguson, T.J. (1979). Accuracy, confidence, and juror perceptions in eyewitness identification. *J Appl Psychol, 64*(4), 440-8.

Wells, G.L., & Loftus, E.F. (2002). 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.

Wells, G.L., & Loftus, E.F. (2003). Eyewitness memory for people and events. In A.M. Goldstein (Ed.), *Handbook of psychology: Forensic psychology, Vol. 11.* (pp. 149-160): John Wiley & Sons, Inc.

Wells, G.L., Malpass, R.S., Lindsay, R.C., Fisher, R.P., Turtle, J.W., & Fulero, S.M. (2000). From the lab to the police station. A successful application of eyewitness research. *Am Psychol, 55*(6), 581-98.

Wells, G.L., & Murray, D.M. (1983). What can psychology say about the Neil v. Biggers criteria for judging eyewitness accuracy? *Journal of Applied Psychology, 68*(3), 347.

Wells, G.L., & Olsen, E.A. (2003). Eyewitness testimony. *Annual Review of Psychology, 54*, 277-295.

Wells, G.L., & Seelau, E.P. (1995). Eyewitness identification: Psychological research and legal policy on lineups. *Psychology, Public Policy, and Law, 1*(4), 765-791.

Wells, G.L., Small, M., Penrod, S.J., Malpass, R.S., Fulero, S.M., & Brimacombe, C.A.E. (1998). Eyewitness identification procedures: Recommendations for lineups and photospreads. *Law and Human Behavior, 22*, 603-647.

Wells, G.L., & Turtle, J.W. (1986). Eyewitness identification: The importance of lineup models. *Psychological Bulletin, 99*(3), 320-329.

Wells, G.L., & Turtle, J.W. (1987). Eyewitness testimony research: Current knowledge and emergent controversies. *Canadian Journal of Behavioural Science, 19*(4), 363-388.

Wickham, L.H., & Morris, P.E. (2003). Attractiveness, distinctiveness, and recognition of faces: attractive faces can be typical or distinctive but are not better recognized. *Am J Psychol, 116*(3), 455-68.

Wilding, E.L., & Rugg, M.D. (1996). An event-related potential study of recognition memory with and without retrieval of source. *Brain: A Journal of Neurology, 119*(3), 889-905.

Wixted, J.T., & Ebbesen, E.B. (1991). On the form of forgetting. *Psychological Science, 2*(6), 409.

Wixted, J.T., & Ebbesen, E.B. (1997). Genuine power curves in forgetting: A quantitative analysis of individual subject forgetting functions. *Memory & Cognition, 25*(5), 731.

Woocher, F.D. (1977). *Storage and retrieval of information from artificial linear orderings.*, Univ Microfilms International.

Wright, D.B. (1993). Misinformation and warnings in eyewitness testimony: a new testing procedure to differentiate explanations. *Memory, 1*(2), 153-66.

Wright, D.B., & McDaid, A.T. (1996). Comparing system and estimator variables using data from real line-ups. *Applied Cognitive Psychology, 10*(1), 75-84.

Wright, D.B., & Stroud, J.N. (2002). Age differences in lineup identification accuracy: people are better with their own age. *Law Hum Behav, 26*(6), 641-54.

Yarmey, A.D. (2003). Eyewitness Identification: Guidelines and Recommendations for Identification Procedures in the United States and in Canada. *Canadian Psychology, 44*(3), 181-189.

Yarmey, A.D. (2004). Eyewitness recall and photo identification: a field experiment. *Psychology, Crime & Law, 10*(1), 53-68.

Yonelinas, A.P. (1994). Receiver-operating characteristics in recognition memory: evidence for a dual-process model. *J Exp Psychol Learn Mem Cogn, 20*(6), 1341-54.

Yonelinas, A.P. (1999). The contribution of recollection and familiarity to recognition and source-memory judgments: a formal dual-process model and an analysis of receiver operating characteristics. *J Exp Psychol Learn Mem Cogn, 25*(6), 1415-34.

Yonelinas, A.P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language, 46*(3), 441.

Yonelinas, A.P., Dobbins, I., Szymanski, M.D., Dhaliwal, H.S., & King, L. (1996). Signal-detection, threshold, and dual-process models of recognition memory: ROCs and conscious recollection. *Conscious Cogn, 5*(4), 418-41.

Appendix A: Tables

Table 1. Experiment 1. Results from the fixed-interval and criminal-based RM ANOVA analyses.

		Fixed-Interval		Criminal-based	
Effects	df	MSE	F	MSE	F
CONDITION (C)	(1,19)	83.30	25.26***	152.89	22.23***
SITE (S)	(5,95)	14.10	12.49***	17.87	13.80***
PHOTO (P)	(5,95)	26.56	36.40***	34.04	44.75***
C x S	(5,95)	1.10	3.19*	2.54	3.71**
C x P	(5,95)	29.95	4.03**	40.20	3.68**
S x P	(25,475)	0.43	5.30***	0.72	5.26***
C x S x P	(25, 475)	0.47	2.13*	0.75	2.63**

(*p < 0.05, **p < 0.01 and ***p < 0.001 after Greenhouse-Geisser correction).

Table 2. Experiment 1. The lineup member associated with the highest z-score and mean z-score for each participant in the standard condition for the fixed interval analysis. ID = identification decision. Bolded items represent z-scores and mean z-scores beyond the statistical cut-off.

Fixed-Interval								
Subject	**Cp3**	**Cpz**	**Cp4**	**P3**	**Pz**	**P4**	**Mean**	**ID**
ERP Line-up A								
01	**F4**	F4	F4	**F4**	C	**C**	F4	FR
03	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
05	C	C	C	F5	C	F5	C	FR
08	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
10	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
12	F3	C	**F1**	C	C	**F1**	C	FR
14	**C**	**C**	**C**	C	**C**	**C**	**C**	CI
16	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
18	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
20	**C**	**C**	**C**	**C**	C	C	**C**	CI
$Total_A$	8	9	8	8	10	8	9	7
ERP Line-up B								
02	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
04	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
06	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
09	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
11	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
15	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
17	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
19	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
21	**C**	C	C	**C**	**C**	C	**C**	CI
22	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
$Total_B$	10	10	10	10	10	10	10	10
$Total_{A+B}$	**18**	**19**	**18**	**18**	**20**	**18**	**19**	**17**
$\%_{A+B}$	**90**	**95**	**90**	**90**	**100**	**90**	**95**	**85**

Correct Identifications (CI)	17/20 (85%)
Misidentifications (MI)	0/20 (0%)
False Rejections (FR)	3/20 (15%)

Table 3. Experiment 1. The lineup member associated with the highest z-score and mean z-score for each participant in the standard condition for the criminal-based analysis. ID = identification decision. Bolded items represent z-scores and mean z-scores beyond the statistical cut-off.

	Criminal-based							
Subject	**Cp3**	**Cpz**	**Cp4**	**P3**	**Pz**	**P4**	**Mean**	**ID**
ERP Line-up A								
01	**F4**	C	C	**F4**	C	**C**	F4	FR
03	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
05	C	C	**C**	F5	C	F5	C	FR
08	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
10	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
12	**C**	C	F1	**C**	C	F1	C	FR
14	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
16	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
18	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
20	**C**	**C**	**C**	**C**	C	C	**C**	CI
Total$_{A}$	9	10	9	8	10	8	**9**	7
ERP Line-up B								
02	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
04	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
06	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
09	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
11	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
15	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
17	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
19	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
21	**C**	C	F4	**C**	**C**	C	**C**	CI
22	**C**	**C**	**C**	C	**C**	**C**	**C**	CI
Total$_{B}$	10	10	9	10	10	10	10	10
Total$_{A+B}$	**19**	**20**	**18**	**18**	**20**	**18**	**19**	**17**
%$_{A+B}$	**95**	**100**	**90**	**90**	**100**	**90**	**95**	**85%**
Correct Identifications (CI)								17/20 (85%)
Misidentifications (MI)								0/20 (0%)
False Rejections (FR)								3/20 (15%)

Table 4. Experiment 1. The lineup member associated with the highest z-score and mean z-score for each participant in the deception condition for the fixed interval analysis. ID = identification decision. Bolded items represent z-scores and mean z-scores beyond the statistical cut-off.

				Fixed-Interval				
Subject	**Cp3**	**Cpz**	**Cp4**	**P3**	**Pz**	**P4**	**Mean**	**ID**
ERP Line-up A								
02	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
04	**F1**	**F1**	**F1**	**F1**	**F1**	**F1**	**F1**	MI
06	F1	C	F1	F1	F1	F1	F1	FR
09	**C**	**C**	C	**C**	**C**	**F1**	**C**	CI
11	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
15	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
17	F1	F1	F1	F4	F1	F1	F1	FR
19	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
21	**F2**	**C**	**F2**	C	C	F2	C	FR
22	**F2**	**C**	C	**F4**	**F2**	**F4**	**F2**	MI
$Total_A$	5	8	6	6	6	4	6	5
ERP Line-up B								
01	**C**	F4	F2	**C**	F4	F4	C	FR
03	C	F1	F1	F1	F1	F1	F1	FR
05	**F1**	**F1**	**C**	**F1**	**F1**	**F1**	**F1**	MI
08	C	**C**	C	**C**	**C**	**C**	**C**	CI
10	C	**C**	**C**	**C**	**C**	**C**	**C**	CI
12	F1	C	C	**C**	**C**	C	C	FR
14	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
16	**F3**	**F3**	**F3**	**F3**	**F3**	**F3**	**F3**	MI
18	C	**C**	**C**	**C**	**C**	**C**	**C**	CI
20	F4	**C**	**C**	**C**	**C**	**C**	**C**	CI
$Total_B$	6	6	7	7	6	6	7	5
$Total_{A+B}$	**11**	**14**	**13**	**13**	**12**	**10**	**13**	**10**
$\%_{A+B}$	**55**	**70**	**65**	**65**	**60**	**50**	**65**	**50**
Correct Identifications (CI)								10/20 (50%)
Misidentifications (MI)								4/20 (20%)
False Rejections (FR)								6/20 (30%)

Table 5. Experiment 1. The lineup member associated with the highest z-score and mean z-score for each participant in the deception condition for the criminal-based analysis. ID = identification decision. Bolded items represent z-scores and mean z-scores beyond the statistical cut-off.

Criminal-based								
Subject	**Cp3**	**Cpz**	**Cp4**	**P3**	**Pz**	**P4**	**Mean**	**ID**
ERP Line-up A								
02	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
04	**F2**	F2	F2	F1	F1	F1	F2	FR
06	F1	C	F1	F1	F1	F1	F1	FR
09	**C**	**C**	C	**C**	**C**	**F1**	**C**	CI
11	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
15	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
17	F4	F4	F1	F4	F4	F1	F4	FR
19	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
21	F2	**C**	F1	**C**	**C**	F2	C	FR
22	**C**	**C**	C	**F4**	F2	**F4**	C	FR
Total$_A$	6	8	6	6	6	4	7	5
ERP Line-up B								
01	**C**	C	**F2**	**C**	C	**F2**	C	FR
03	C	F1	C	**C**	F1	**C**	C	FR
05	F1	C	C	**F1**	C	F1	F1	FR
08	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
10	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
12	C	C	**F1**	**C**	**C**	**F1**	**C**	CI
14	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
16	**F3**	**F3**	**F3**	**F3**	**F3**	**F3**	**F3**	MI
18	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
20	F4	**C**	**C**	C	**C**	**C**	**C**	CI
Total$_B$	7	8	7	8	8	6	**8**	6
Total$_{A+B}$	13	16	13	14	14	10	**15**	**11**
%$_{A+B}$	65	80	65	70	70	50	**75**	**55%**

Correct Identifications (CI)	11/20 (55%)
Misidentifications (MI)	1/20 (5%)
False Rejections (FR)	8/20 (40%)

Table 6. Experiment 1. Certainty ratings for the standard condition.

	Certainty Ratings						
Subject #	**F1**	**F2**	**F3**	**F4**	**F5**	**C**	**ID**
ERP-lineup A							
01	4	6	0	4	0	**9**	CI
03	0	0	0	0	0	**10**	CI
05	0	3	0	0	0	**9**	CI
08	1	0	0	0	0	**10**	CI
10	0	0	0	0	0	**10**	CI
12	0	**5**	0	0	0	**5**	CI
14	0	0	0	5	0	**10**	CI
16	0	0	0	4	0	**9**	CI
18	0	0	0	0	0	**10**	CI
20	0	0	0	0	0	**10**	CI
M_A	0.5	1.4	0	1.3	0	**9.2**	10
ERP-up B							
02	0	0	0	0	0	**10**	CI
04	0	0	0	0	0	**10**	CI
06	0	0	0	0	0	**10**	CI
09	0	0	0	0	0	**10**	CI
11	0	**5**	0	0	0	**5**	CI
15	0	0	0	0	0	**10**	CI
17	0	0	0	0	0	**10**	CI
19	0	0	0	0	0	**10**	CI
21	0	0	0	0	0	**10**	CI
22	0	2	3	0	0	**10**	CI
M_B	0	0.7	0.3	0	0	**9.5**	10
$\mathbf{M_{A+B}}$	0.3	1.1	0.2	0.7	0	**9.4**	20
Correct Identifications (CI)							20/20 (100%)
Misidentifications (MI)							0/20 (100%)
False Rejections (FR)							0/20 (100%)

Table 7. Experiment 1. Certainty ratings for the deception condition.

	Certainty Ratings						
Subject #	**F1**	**F2**	**F3**	**F4**	**F5**	**C**	**ID**
ERP-lineup A							
02	0	0	0	0	0	**8**	CI
04	2	5	0	0	0	**9**	CI
06	0	0	0	0	0	**9**	CI
09	0	0	0	0	0	**10**	CI
11	1	0	0	0	0	**9**	CI
15	0	0	0	0	0	**10**	CI
17	3	4	1	6	0	**9**	CI
19	0	3	0	3	0	**7.5**	CI
21	0	2	0	0	0	**10**	CI
22	1	0	0	0	0	**10**	CI
M_A	0.7	1.4	0.1	0.9	0	**9.2**	10
ERP-lineup B							
01	0	0	0	0	0	**10**	CI
03	0	5	5	0	0	**5**	CI
05	0	0	0	0	0	**10**	CI
08	0	0	0	0	0	**10**	CI
10	0	0	0	0	0	**10**	CI
12	0	0	0	0	0	**10**	CI
14	0	0	0	0	0	**10**	CI
16	0	0	0	0	0	**10**	CI
18	0	0	0	0	0	**10**	CI
20	0	0	0	0	0	**10**	CI
M_B	0	0.5	0.5	0	0	**9.5**	CI
$\boldsymbol{M_{A+B}}$	**0.4**	**1.0**	**0.3**	**0.5**	**0**	9.3	10
Correct Identifications (CI)						20/20 (100%)	
Misidentifications (MI)						0/20 (0%)	
False Rejections (FR)						0/20 (100%)	

Table 8. Experiment 1. Percentage of time participants classified each lineup member as the criminal in the standard condition. ID = identification decision.

% Button Press Responses							
Subject #:	**F1**	**F2**	**F3**	**F4**	**F5**	**C**	**ID**
ERP-lineup A							
01	0	0	0	0	0	**92.5**	CI
03	0	0	0	0	0	**95**	CI
05	0	2.5	0	0	0	**97.5**	CI
08	0	0	0	0	0	**97.5**	CI
10	0	0	0	0	0	**100**	CI
12	0	0	0	0	0	**97.5**	CI
14	0	0	0	0	0	**100**	CI
16	0	0	0	0	0	**97.5**	CI
18	0	5	0	0	0	**97.5**	CI
20	0	0	0	0	0	**97.5**	CI
M_A	0	0.8	0	0	0	**97.3**	10
ERP-line up B							
02	0	0	0	0	0	**100**	CI
04	0	0	0	0	0	**100**	CI
06	0	0	0	0	0	**95**	CI
09	0	0	0	0	0	**95**	CI
11	0	0	0	0	0	**97.5**	CI
15	0	0	0	0	0	**97.5**	CI
17	0	0	0	0	0	**80**	CI
19	0	0	0	0	0	**92.5**	CI
21	0	0	0	0	0	**92.5**	CI
22	0	0	0	0	0	**100**	CI
M_B	0	0	0	0	0	**95**	10
$\boldsymbol{M_{A+B}}$	0	0.4	0	0	0	**96.1**	20
Correct Identifications (CI)							20/20 (100%)
Misidentifications (MI)							0/20 (0%)
False Rejections (FR)							0/20 (0%)

Table 9. Experiment 1. Percentage of times each participants classified each lineup member as the criminal in the deception condition. ID = identification decision.

% Button Press Responses							
Subject #:	**F1**	**F2**	**F3**	**F4**	**F5**	**C**	**ID**
ERP-lineup A							
02	0	0	0	0	0	0	FR
04	0	0	0	0	0	0	FR
06	0	0	0	0	0	0	FR
09	0	0	0	0	0	0	FR
11	0	0	0	0	0	0	FR
15	0	0	0	0	0	0	FR
17	0	0	0	0	0	0	FR
19	0	0	0	0	0	0	FR
21	0	0	0	0	0	0	FR
22	0	0	0	0	0	0	FR
M_A	0	0	0	0	0	0	0
ERP-line up B							
01	0	0	0	0	0	0	FR
03	0	0	0	0	0	0	FR
05	0	0	0	0	0	0	FR
08	0	0	0	0	0	0	FR
10	0	0	0	0	0	0	FR
12	0	0	0	0	0	0	FR
14	0	0	0	0	0	0	FR
16	0	0	0	0	0	0	FR
18	0	0	0	0	0	0	FR
20	0	0	0	0	0	0	FR
M_B	0	0	0	0	0	0	0
M_{A+B}	0	0	0	0	0	0	0
Correct Identifications (CI)							0/20 (100%)
Misidentifications (MI)							0/20 (0%)
False Rejections (FR)							20/20 (100%)

Table 10. Experiment 1. Summary chart of identification decisions for the standard condition.

Identification Decisions			
Subject #:	**Certainty Ratings**	**Button Press**	**ERP Patterns**
ERP-lineup A			
01	CI	CI	FR
03	CI	CI	CI
05	CI	CI	FR
08	CI	CI	CI
10	CI	CI	CI
12	CI	CI	FR
14	CI	CI	CI
16	CI	CI	CI
18	CI	CI	CI
20	CI	CI	CI
$Total_A$	10	10	7
ERP-line up B			
02	CI	CI	CI
04	CI	CI	CI
06	CI	CI	CI
09	CI	CI	CI
11	CI	CI	CI
15	CI	CI	CI
17	CI	CI	CI
19	CI	CI	CI
21	CI	CI	CI
22	CI	CI	CI
$Total_B$	10	10	10
Correct Identifications (CI)	20 (100%)	20 (100%)	17/20 (85%)
Misidentifications (MI)	0 (0%)	0 (0%)	0 (0%)
False Rejections (FR)	0 (0%)	0 (0%)	3/20 (15%)

Table 11. Experiment 1. Summary chart of identification decisions for the deception condition.

Identification Decisions			
Subject #:	**Certainty Ratings**	**Button Press**	**ERP Patterns**
ERP-lineup A			
02	CI	FR	CI
04	CI	FR	FR
06	CI	FR	FR
09	CI	FR	CI
11	CI	FR	CI
15	CI	FR	CI
17	CI	FR	FR
19	CI	FR	CI
21	CI	FR	FR
22	CI	FR	FR
$Total_A$	10	0	5
ERP-line up B			
01	CI	FR	FR
03	CI	FR	FR
05	CI	FR	FR
08	CI	FR	CI
10	CI	FR	CI
12	CI	FR	CI
14	CI	FR	CI
16	CI	FR	MI
18	CI	FR	CI
20	CI	FR	CI
$Total_B$	10	0	6
Correct Identifications (CI)	20 (100%)	20 (100%)	11/20 (55%)
Misidentifications (MI)	0 (0%)	0 (0%)	1/20 (5%)
False Rejections (FR)	0 (0%)	0 (0%)	8/20 (40%)

Table 12. Experiment 2a. Results from the fixed-interval and criminal-based RM ANOVA analyses.

Effects	df	Fixed-Interval MSE	Fixed-Interval F	Criminal-based MSE	Criminal-based F
CONDITION(C)	(2,46)	115.07	2.85	153.55	2.43
SITE (S)	(5,115)	18.79	12.47***	24.02	14.82***
PHOTO (P)	(5,115)	64.43	39.84***	84.55	45.89***
C x S	(10,230)	1.84	0.89	4.67	0.97
C x P	(10,230)	37.03	1.15	45.61	0.98
S x P	(25,575)	0.80	8.11***	1.40	7.22***
C x S x P	(50, 1150)	0.64	0.97	1.03	1.25

(*$p < 0.05$, **$p < 0.01$ and ***$p < 0.001$ after Greenhouse-Geisser correction).

Table 13. Experiment 2a. The lineup member associated with the highest z-score and mean z-score for each participant in the no-delay condition for the fixed-interval analysis. ID = identification decision. Bolded items represent z-scores and mean z-scores beyond the statistical cut-off.

Fixed-Interval								
Subject	**Cp3**	**Cpz**	**Cp4**	**P3**	**Pz**	**P4**	**Mean**	**ID**
ERP Line-up A								
02	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
06	F3	**C**	**C**	**F3**	**C**	**C**	**C**	CI
12	**F1**	**F1**	C	**F1**	C	C	F1	FR
20	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
22	**C**	**C**	**C**	**C**	**C**	**F1**	**C**	CI
29	F3	F3	**F3**	F3	F3	F5	F3	FR
Total$_A$	3	4	5	3	5	4	4	4
ERP Line-up B								
01	**F1**	**F1**	**F1**	**F1**	**F1**	F1	**F1**	MI
05	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
11	F1	**C**	**F1**	F3	**C**	F1	**C**	CI
15	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
23	**F4**	**F4**	**F4**	**F4**	**F4**	**F4**	**F4**	MI
24	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
Total$_B$	3	4	3	3	4	3	4	4
ERP Line-up C								
04	**F4**	**F4**	**F4**	**C**	**F4**	**F4**	**F4**	MI
10	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
14	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
18	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
21	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
28	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
Total$_C$	5	5	5	6	5	5	5	5
ERP line-up D								
07	**F1**	**F1**	**F1**	**F1**	**F1**	**F1**	**F1**	MI
09	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
13	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
25	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
26	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
27	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
Total$_D$	5	5	5	5	5	5	5	5
Total$_{ABCD}$	**16**	**18**	**18**	**17**	**19**	**17**	**18**	**18**
%$_{ABCD}$	**66**	**75**	**75**	**71**	**79**	**71**	**75**	**75**
Correct Identification (CI)								18/24 (75%)
Misidentifications (MI)								4/24 (17%)
False Rejection (FR)								2/24 (8%)

Table 14. Experiment 2a. The lineup member associated with the highest z-score and mean z-score for each participant in the no-delay condition for the criminal-based analysis. ID = identification decision. Bolded items represent z-scores and mean z-scores beyond the statistical cut-off.

Criminal-based								
Subject	**Cp3**	**Cpz**	**Cp4**	**P3**	**Pz**	**P4**	**Mean**	**ID**
ERP Line-up A								
02	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
06	F3	**C**	**C**	**F3**	**C**	**C**	**C**	CI
12	**F1**	**F1**	C	**F1**	**F1**	C	F1	FR
20	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
22	**C**	**C**	F1	C	**C**	**F1**	**C**	CI
29	**C**	**C**	F3	C	**C**	F3	**C**	CI
$Total_A$	4	5	4	4	5	4	5	5
ERP Line-up B								
01	**F1**	**F1**	**F1**	**F1**	**F1**	**F1**	**F1**	MI
05	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
11	F1	C	**F1**	**F3**	**C**	F1	C	FR
15	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
23	**F4**	**F4**	**F4**	**F4**	**F4**	**F4**	**F4**	MI
24	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
$Total_B$	3	4	3	3	4	3	4	3
ERP Line-up C								
04	**C**	C	C	**C**	F4	F4	**C**	CI
10	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
14	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
18	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
21	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
28	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
$Total_C$	6	6	6	6	5	5	6	6
ERP line-up D								
07	C	C	**C**	C	C	**C**	C	FR
09	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
13	**C**	**C**	**C**	**C**	C	**C**	**C**	CI
25	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
26	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
27	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
$Total_D$	6	6	6	6	6	6	6	5
$Total_{ABCD}$	**19**	**21**	**19**	**19**	**20**	**18**	**21**	**19**
$\%_{ABCD}$	**79**	**88**	**79**	**79**	**83**	**75**	**88**	**79**
Correct Identification (CI)								19/24 (79%)
Misidentifications (MI)								2/24 (8%)
False Rejection (FR)								3/24 (13%)

Table 15. Experiment 2a. The lineup member associated with the highest z-score and mean z-score for each participant in the 1-hour delay condition for the fixed-interval analysis. ID = identification decision. Bolded items represent z-scores and mean z-scores beyond the statistical cut-off.

Fixed-Interval								
Subject	**Cp3**	**Cpz**	**Cp4**	**P3**	**Pz**	**P4**	**Mean**	**ID**
ERP Line-up A								
01	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
05	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
09	**C**	**C**	C	**C**	**C**	**C**	**C**	CI
13	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
26	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
27	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
$Total_A$	6	6	6	6	6	6	6	6
ERP Line-up B								
02	**C**	**C**	**C**	F4	**C**	**C**	**C**	CI
06	**F3**	**F3**	**F3**	**F3**	**F3**	**F3**	**F3**	MI
10	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
14	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
18	F3	F3	C/F3	F3	F1	C	F3	FR
28	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
$Total_B$	4	4	5	3	4	5	4	4
ERP Line-up C								
07	**C**	**C**	C	**C**	**C**	C	**C**	CI
11	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
15	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
23	**F1**	**F1**	F1	**F1**	**F1**	F4	**F1**	MI
24	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
25	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
$Total_C$	5	5	5	5	5	5	5	5
ERP line-up D								
04	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
12	**F3**	F2	F2	F5	F5	F5	F5	FR
20	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
21	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
22	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
29	**C**	**C**	C	**C**	**C**	**C**	**C**	CI
$Total_D$	5	5	5	5	5	5	5	5
$Total_{ABCD}$	**20**	**20**	**21**	**19**	**20**	**21**	**20**	**20**
$\%_{ABCD}$	**83**	**83**	**88**	**79**	**83**	**88**	**83**	**83**
Correct Identifications (CI)								20/24 (83%)
Misidentifications (MI)								2/24 (8%)
False Rejections (FR)								2/24 (8%)

Table 16. Experiment 2a. The lineup member associated with the highest z-score and mean z-score for each participant in the 1-hour delay condition for the criminal-based analysis. ID = identification decision. Bolded items represent z-scores and mean z-scores beyond the statistical cut-off.

Criminal-based								
Subject	**Cp3**	**Cpz**	**Cp4**	**P3**	**Pz**	**P4**	**Mean**	**ID**
ERP Line-up A								
01	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
05	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
09	**C**	**C**	**F1**	**C**	**C**	**C**	**C**	CI
13	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
26	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
27	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
$Total_A$	6	6	5	6	6	6	6	6
ERP Line-up B								
02	**C**	**C**	C	**F4**	**C**	**C**	**C**	CI
06	**F3**	**F3**	**F3**	**F3**	**F3**	**F5**	**F3**	MI
10	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
14	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
18	**C**	**C**	**C**	**C**	**C**	C	**C**	CI
28	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
$Total_B$	5	5	5	4	5	5	5	5
ERP Line-up C								
07	C	C	C	F5	C	C	C	FR
11	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
15	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
23	**C**	**F1**	C	**F1**	F1	**F4**	F1	FR
24	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
25	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
$Total_C$	6	5	6	4	5	5	5	4
ERP line-up D								
04	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
12	**F3**	F2	F2	F3	F2	F2	F3	FR
20	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
21	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
22	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
29	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
$Total_D$	5	5	5	5	5	5	5	5
$Total_{ABCD}$	**22**	**21**	**21**	**19**	**21**	**21**	**21**	20
$\%_{ABCD}$	**92**	**88**	**88**	**79**	**88**	**88**	**88**	**83**
Correct Identifications (CI)								20/24 (83%)
Misidentifications (MI)								1/24 (4%)
False Rejections (FR)								3/24 (13%)

Table 17. Experiment 2a. The lineup member associated with the highest z-score and mean z-score for each participant in the 1-week delay condition for the fixed-interval analysis. ID = identification decision. Bolded items represent z-scores and mean z-scores beyond the statistical cut-off.

Fixed-Interval								
Subject	**Cp3**	**Cpz**	**Cp4**	**P3**	**Pz**	**P4**	**Mean**	**ID**
ERP Line-up A								
07	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
11	**F3**	**F3**	**F3**	**F3**	**F3**	**F3**	**F3**	MI
15	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
23	**F4**	**F4**	**F4**	**F4**	**F4**	**F4**	**F4**	MI
24	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
25	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
$Total_A$	4	4	4	4	4	4	4	4
ERP Line-up B								
04	F5	F5	F2	F5	**F5**	**F2**	F5	FR
12	**F5**	**F5**	**F5**	**F5**	**F5**	**F5**	**F5**	MI
20	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
21	**F5**	**F5**	**F5**	**F5**	**F5**	**F5**	**F5**	MI
22	**F5**	**F5**	**F5**	**F5**	**F5**	**F5**	**F5**	MI
29	**F3**	F3	F3	F3	**F3**	**F3**	**F3**	MI
$Total_B$	1	1	1	1	1	1	1	1
ERP Line-up C								
01	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
05	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
09	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
13	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
26	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
27	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
$Total_C$	6	6	6	6	6	6	6	6
ERP line-up D								
02	**C**	**C**	**F5**	**C**	**C**	**C**	**C**	CI
06	**F4**	**F5**	**F5**	**F4**	**F5**	F4	**F5**	MI
10	C	F2	F2	F4	F4	**F4**	F4	FR
14	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
18	**F4**	**C**	**C**	**F4**	**C**	F5	F4	FR
28	F5	**F5**	**F5**	F5	F5	F5	F5	FR
$Total_D$	3	3	2	2	3	2	2	2
$Total_{ABCD}$	**14**	**14**	**13**	**13**	**14**	**13**	**13**	**13**
$\%_{ABCD}$	**58**	**58**	**54**	**54**	**58**	**54**	**54**	**54**
Correct Identifications (CI)								13/24 (54%)
Misidentifications (MI)								7/24 (29%)
False Rejections (FR)								4/24 (17%)

Table 18. Experiment 2a. The lineup member associated with the highest z-score and mean z-score for each participant in the 1-week delay condition for the criminal-based analysis. ID = identification decision. Bolded items represent z-scores and mean z-scores beyond the statistical cut-off.

Criminal-based								
Subject	**Cp3**	**Cpz**	**Cp4**	**P3**	**Pz**	**P4**	**Mean**	**ID**
ERP Line-up A								
07	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
11	**F3**	**F3**	**F3**	**F3**	**F3**	**F3**	**F3**	MI
15	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
23	**F4**	**F4**	**F4**	**F4**	**F4**	**F4**	**F4**	MI
24	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
25	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
$Total_A$	4	4	4	4	4	4	4	4
ERP Line-up B								
04	F2	F5	**F2**	**F2**	F2	**F2**	**F2**	MI
12	**F5**	**F5**	**F5**	**F5**	**F5**	**F5**	**F5**	MI
20	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
21	**F5**	**F5**	**F5**	**F5**	**F5**	**F2**	**F5**	MI
22	**F5**	**F5**	**F5**	**F5**	**F5**	**F5**	**F5**	MI
29	F1	**F3**	**F3**	**F3**	**F3**	**F3**	**F3**	MI
$Total_B$	1	1	1	1	1	1	1	1
ERP Line-up C								
01	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
05	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
09	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
13	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
26	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
27	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
$Total_C$	6	6	6	6	6	6	6	6
ERP line-up D								
02	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
06	**F4**	**F5**	**F5**	**F4**	F4	F2	F4	FR
10	C	F2	F2	F4	F4	**F4**	F4	FR
14	**C**	**C**	**C**	**C**	**C**	**C**	**C**	CI
18	C	**C**	**C**	F4	**C**	C	**C**	CI
28	F3	F5	F3	**F5**	**F5**	F5	F5	FR
$Total_D$	4	3	3	2	3	3	3	3
$Total_{ABCD}$	**15**	**14**	**14**	**13**	**14**	**14**	**14**	**14**
$\%_{ABCD}$	**63**	**58**	**58**	**54**	**58**	**58**	**58**	**58**
Correct Identifications (CI)								14/24 (58%)
Misidentifications (MI)								7/24 (29%)
False Rejections (FR)								3/24 (13%)

Table 19. Experiment 2a. Certainty ratings for the no-delay condition.

	Certainty Ratings						
Subject #	**F1**	**F2**	**F3**	**F4**	**F5**	**C**	**ID**
ERP Line-up A							
02	0	4	0	0	3	**8**	CI
06	0	0	0	0	0	3	FR
12	0	0	0	0	0	**9**	CI
20	0	0	0	0	3	**7**	CI
22	0	0	0	0	0	**9**	CI
29	0	1	0	0	1	3	FR
M_A	0.0	0.8	0.0	0.0	1.2	6.5	4
ERP Line-up B							
01	**8**	0	0	0	0	5	MI
05	0	0	1	1	0	**8**	CI
11	0	4	0	**5**	0	0	MI
15	0	0	0	0	1	**8**	CI
23	1	0	0	0	1	1	FR
24	0	0	0	0	0	**8**	CI
M_B	1.5	0.7	0.2	1.0	0.3	5.0	3
ERP Line-up C							
04	0	0	0	0	0	**6**	CI
10	0	0	0	0	0	**5**	CI
14	0	0	0	0	0	**9**	CI
18	2	0	1	2	1	**9**	CI
21	0	0	0	0	0	**10**	CI
28	0	0	0	0	0	**10**	CI
M_C	0.3	0.0	0.2	0.3	0.2	8.2	6
ERP Line-up D							
07	5	2	3	0	5	**6**	CI
09	0	0	0	0	0	**5**	CI
13	0	0	0	0	0	1	FR
25	0	1	1	2	2	**6**	CI
26	0	0	0	0	0	**5**	CI
27	0	0	0	0	0	**6**	CI
M_D	0.8	0.5	0.7	0.3	1.2	4.8	5
$\mathbf{M_{ABCD}}$	**0.7**	**0.5**	**0.3**	**0.4**	**0.7**	**6.1**	**18**
Correct Identifications (CI)						18/24 (75%)	
Misidentifications (MI)						2/24 (8%)	
False Rejections (FR)						4/24 (17%)	

Table 20. Experiment 2a. Certainty ratings for the 1-hour delay condition.

Certainty Ratings							
Subject	**F1**	**F2**	**F3**	**F4**	**F5**	**C**	**ID**
ERP Line-up A							
01	0	0	0	0	0	**10**	CI
05	2	0	0	1	1	**10**	CI
09	0	0	0	0	0	**10**	CI
13	0	0	0	0	0	**9**	CI
26	0	0	0	0	0	**10**	CI
27	0	0	0	0	0	**10**	CI
M_A	0.3	0.0	0.0	0.2	0.2	9.8	6
ERP Line-up B							
02	5	5	0	0	4	**6**	CI
06	0	0	0	0	3	**7**	CI
10	0	0	0	0	0	**10**	CI
14	0	0	0	0	0	**10**	CI
18	2	3	2	0	3	3	FR
28	0	0	0	0	0	**10**	CI
M_B	1.2	1.3	0.3	0.0	1.7	7.7	5
ERP Line-up C							
07	0	0	2	2	0	**9**	CI
11	0	0	0	0	0	**10**	CI
15	0	0	0	0	0	**10**	CI
23	1	0	0	1	3	**5**	CI
24	0	0	0	0	0	**10**	CI
25	1	1	0	4	0	**9**	CI
M_C	0.3	0.2	0.3	1.2	0.5	8.8	6
ERP Line-up D							
04	0	0	0	0	0	**8**	CI
12	1	**6**	0	0	0	2	MI
20	2	4	2	2	2	**8**	CI
21	0	0	0	0	0	**10**	CI
22	4	0	0	0	0	**9**	CI
29	4	0	0	0	1	**9**	CI
M_D	1.8	1.7	0.3	0.3	0.5	7.7	5
M_{ABCD}	**0.9**	**0.8**	**0.3**	**0.4**	**0.7**	**8.5**	**22**
Correct Identifications (CI)							22/24 (92%)
Misidentifications (MI)							1/24 (4%)
False Rejections (FR)							1/24 (4%)

Table 21. Experiment 2a. Certainty ratings for the 1-week delay condition.

	Certainty Ratings						
Subject	**F1**	**F2**	**F3**	**F4**	**F5**	**C**	**ID**
ERP Line-up A							
07	3	4	2	0	4	**8**	CI
11	0	2	3	2	2	**5**	CI
15	0	0	0	0	0	**10**	CI
23	0	0	0	**5**	4	2	MI
24	0	0	0	0	0	**10**	CI
25	1	3	1	1	2	**5**	CI
M_A	0.7	1.5	1.0	1.3	2.0	6.7	5
ERP Line-up B							
04	0	0	0	0	2	0	FR
12	0	0	0	0	**9**	2	MI
20	2	1	1	1	2	**8**	CI
21	0	0	3	0	**8**	0	MI
22	0	0	0	0	**8**	5	MI
29	3	0	**5**	0	0	2	MI
M_B	0.8	0.2	1.5	0.2	4.8	2.8	1
ERP Line-up C							
01	0	0	0	0	0	**10**	CI
05	4	1	0	0	0	**7**	CI
09	0	0	0	0	0	**10**	CI
13	0	0	0	0	0	**8**	CI
26	0	0	0	0	0	**10**	CI
27	0	0	0	0	0	**9**	CI
M_C	0.7	0.2	0.0	0.0	0.0	9.0	6
ERP Line-up D							
02	1	1	1	0	3	**7**	CI
06	4	0	0	0	**8**	0	MI
10	0	0	**5**	0	0	0	MI
14	0	0	0	0	3	**6**	CI
18	1	1	2	1	**6**	3	MI
28	1	0	0	0	0	0	FR
M_D	1.2	0.3	1.3	0.2	3.3	2.7	2
$\mathbf{M_{ABCD}}$	**0.8**	**0.5**	**1.0**	**0.4**	**2.5**	**5.3**	**14**
Correct Identifications (CI)							14/24 (58%)
Misidentifications (MI)							8/24 (33%)
False Rejections (FR)							2/24 (8%)

Table 22. Experiment 2a. Percentage of times participants classified each lineup member as the criminal in the no-delay condition. * denotes when participants refrained from making a button press response more than 10% of the time. ID = identification decision.

			Button Press				
Subject	**F1**	**F2**	**F3**	**F4**	**F5**	**C**	**ID**
ERP Line-up A							
02	0	0	0	0	0	**92.5**	CI
06	0	0	0	0	0	5	FR
12	0	2.5	0	0	0	**95**	CI
20	0	0	0	0	0*	**90**	CI
22	0	2.5	0	0	0	45*	FR
29	0	0	0	0	0	2.5	FR
M_A	0	0.8	0	0	0	47.5	3
ERP Line-up B							
01	**92.5**	0	0	0	2.5	0	MI
05	0	0	0	2.5	0	**97.5**	CI
11	0	2.5	0	0	0	0	FR
15	0	0	0	0	0	**92.5**	CI
23	0	0	0	0	0	0	FR
24	0	0	0	0	0	**97.5**	CI
M_B	15.4	0.4	0.0	0.4	0.4	47.9	3
ERP Line-up C							
04	0	0	0	0	0	47.5*	FR
10	0	0	2.5	0	0	20	FR
14	0	0	0	0	0	**92.5**	CI
18	0	0	0	0	0	**100**	CI
21	0	0	0	0	0	**100**	CI
28	0	0	0	0	0	**97.5**	CI
M_C	0.0	0.0	0.4	0.0	0.0	**68.3**	4
ERP Line-up D							
07	0	0	0	0	0	0*	FR
09	0	0	0	0	0	0	FR
13	0	0	0	0	0	0	FR
25	0	0	0	0*	0	**87.5**	CI
26	2.5	0	0	0	0	2.5	FR
27	0	0	0	0	0	**95**	CI
M_D	0.4	0.0	0.0	0.0	0.0	30.8	2
M_{ABCD}	4.0	0.3	0.1	0.1	0.1	48.6	12
Correct Identifications (CI)							12/24 (50%)
Misidentifications (MI)							1/24 (4%)
False Rejections (FR)							11/24 (46%)

Table 23. Experiment 2a. Percentage of times participants classified each lineup member as the criminal in the 1-hour delay condition. * denotes when participants refrained from making a button press response more than 10% of the time. ID = identification decision.

	Button Press						
Subject	**S1**	**S2**	**S3**	**S4**	**S5**	**C**	**ID**
ERP Line-up A							
01	0	0	0	0	0	**100**	CI
05	0	0	0	0	0	**100**	CI
09	0	0	0	0	0	**100**	CI
13	0	0	0	0	0	**100**	CI
26	0	0	0	0	0	**92.5**	CI
27	0	0	0	0	0	**100**	CI
M_A	0	0	0	0	0	**98.8**	6
ERP Line-up B							
02	0*	7.5*	0	0	7.5	**85**	CI
06	0	0	**92.5**	0	0	0	MI
10	0	0	0	0	0	**92.5**	CI
14	0	0	0	0	0	**97.5**	CI
18	0	0	0	0	0	15	FR
28	0	0	0	0	0	**100**	CI
M_B	0	1.3	15.4	0	1.3	**65**	4
ERP Line-up C							
07	0	0	0	0	0	**100**	CI
11	0	0	0	0	0	**97.5**	CI
15	0	0	0	0	0	**95**	CI
23	0	0	0	0	0	0	FR
24	0	0	0	0	0	**100**	CI
25	0	0	0	0	0	**97.5**	CI
M_C	0	0	0	0	0	**81.7**	5
ERP Line-up D							
04	2.5	0	0	0	0*	**77.5***	CI
12	0*	0*	0	0	0*	0*	FR
20	0	0	0	0	0	**92.5**	CI
21	0	0	0	0	0	**100**	CI
22	0	0	0	0	0	**97.5**	CI
29	2.5	0	2.5	0	0	**90**	CI
M_D	0.8	0.0	0.4	0.0	0.0	**76.3**	5
M_{ABCD}	0.2	0.3	4.0	0.0	0.3	**80.4**	20
Correct Identifications (CI)							20/24 (83%)
Misidentifications (MI)							1/24 (4%)
False Rejections (FR)							3/24 (13%)

Table 24. Experiment 2a. Percentage of times participants classified each lineup member as the criminal in the 1-week delay condition. * denotes when participants refrained from making a button press response more than 10% of the time. ID = identification decision.

	Button Press						
Subject	**S1**	**S2**	**S3**	**S4**	**S5**	**C**	**ID**
ERP Line-up A							
07	0	0	0	0	0	**97.5**	CI
11	0	0	30*	0	0	15	FR
15	0	0	0	0	0	**97.5**	CI
23	0	0	0	**85**	0	2.5	MI
24	0	0	0	0	0	**97.5**	CI
25	0	0	0	0	0*	**90**	CI
M_A	0	0	0	14.2	0	**66.7**	4
ERP Line-up B							
04	0	0	0	0	0	0	FR
12	2.5	0	0	0	**77.5**	25	MI
20	0	0	0	0	2.5	**100**	CI
21	0	0	0	0	**90**	0	MI
22	0	0	2.5	0	**95**	12.5	MI
29	10	0	37.5*	0	0	0	FR
M_B	2.1	0	0.4	0	44.2	22.9	1
ERP Line-up C							
01	0	0	0	0	0	**100**	CI
05	2.5	0	0	0	0	35*	FR
09	0	0	0	0	0	**95**	CI
13	0	0	0	0	0	**90**	CI
26	0	0	0	2.5	0	**95**	CI
27	0	0	0	0	0	**100**	CI
M_C	0.4	0	0	0.4	0	**80**	5
ERP Line-up D							
02	2.5	2.5	0	0	0	**90**	CI
06	0	0	0	0	**87.5**	0	MI
10	0	0	0	0	0	0	FR
14	0	0	0	0	0	**97.5**	CI
18	0	0	0	0	45	45	FR
28	0	0	0	0	0	0	FR
M_D	0.4	0.4	0	0	22.1	38.8	2
M_{ABCD}	0.7	0.1	0.1	3.6	16.6	**52.1**	12
Correct Identifications (CI)							12/24 (50%)
Misidentifications (MI)							5/24 (21%)
False Rejections (FR)							7/24 (29%)

Table 25. Experiment 2a. Summary chart of identification decisions for the no-delay condition.

Identification Decisions			
Subject #	**Certainty Ratings**	**Button Press**	**ERP Patterns**
ERP Line-up A			
02	CI	CI	CI
06	FR	FR	CI
12	CI	CI	FR
20	CI	CI	CI
22	CI	FR	CI
29	FR	FR	CI
$Total_A$	4	3	5
ERP Line-up B			
01	MI	MI	MI
05	CI	CI	CI
11	MI	FR	FR
15	CI	CI	CI
23	FR	FR	MI
24	CI	CI	CI
$Total_B$	3	3	3
ERP Line-up C			
04	CI	FR	CI
10	CI	FR	CI
14	CI	CI	CI
18	CI	CI	CI
21	CI	CI	CI
28	CI	CI	CI
$Total_C$	6	4	6
ERP Line-up D			
07	CI	FR	FR
09	CI	FR	CI
13	FR	FR	CI
25	CI	CI	CI
26	CI	FR	CI
27	CI	CI	CI
$Total_D$	5	2	5
Correct Identifications (CI)	18 (75%)	12 (50%)	19 (79%)
Misidentifications (MI)	2 (8%)	1 (4%)	2 (8%)
False Rejections (FR)	4 (17%)	11 (46%)	3 (13%)

Table 26. Experiment 2a. Summary chart of identification decisions for the 1-hour delay condition.

Identification Decisions			
Subject #	**Certainty Ratings**	**Button Press**	**ERP Patterns**
ERP Line-up A			
01	CI	CI	CI
05	CI	CI	CI
09	CI	CI	CI
13	CI	CI	CI
26	CI	CI	CI
27	CI	CI	CI
$Total_A$	6	6	6
ERP Line-up B			
02	CI	CI	CI
06	CI	MI	MI
10	CI	CI	CI
14	CI	CI	CI
18	FR	FR	CI
28	CI	CI	CI
$Total_B$	5	4	5
ERP Line-up C			
07	CI	CI	FR
11	CI	CI	CI
15	CI	CI	CI
23	CI	FR	FR
24	CI	CI	CI
25	CI	CI	CI
$Total_C$	6	5	4
ERP Line-up D			
04	CI	CI	CI
12	MI	FR	FR
20	CI	CI	CI
21	CI	CI	CI
22	CI	CI	CI
29	CI	CI	CI
$Total_D$	5	5	5
Correct Identifications (CI)	22 (92%)	20 (83%)	20 (83%)
Misidentifications (MI)	1 (4%)	1 (4%)	1 (4%)
False Rejections (FR)	1 (4%)	3 (13%)	3 (13%)

Table 27. Experiment 2a. Summary chart of identification decisions for the 1-week condition.

Identification Decisions			
Subject #	**Certainty Ratings**	**Button Press**	**ERP Patterns**
ERP Line-up A			
07	CI	CI	CI
11	CI	FR	MI
15	CI	CI	CI
23	MI	MI	MI
24	CI	CI	CI
25	CI	CI	CI
$Total_A$	5	4	4
ERP Line-up B			
04	FR	FR	MI
12	MI	MI	MI
20	CI	CI	CI
21	MI	MI	MI
22	MI	MI	MI
29	MI	FR	MI
$Total_B$	1	1	1
ERP Line-up CI			
01	CI	CI	CI
05	CI	FR	CI
09	CI	CI	CI
13	CI	CI	CI
26	CI	CI	CI
27	CI	CI	CI
$Total_C$	6	5	6
ERP Line-up D			
02	CI	CI	CI
06	MI	MI	FR
10	MI	FR	FR
14	CI	CI	CI
18	MI	FR	CI
28	FR	FR	FR
$Total_D$	2	2	3
Correct Identifications (CI)	14 (58%)	12 (50%)	14 (58%)
Misidentifications (MI)	8 (33%)	5 (21%)	7 (29%)
False Rejections (FR)	2 (8%)	7 (29%)	3 (13%)

Table 28. Experiment 2b. Results from the RM ANOVA analyses (N=24) for the CA condition alone.

	Fixed-interval			IS-based	
Effects	df	MSE	F	MSE	F
SITE (S)	(5,115)	7.90	13.01***	12.43	16.81***
PHOTO (P)	(5,115)	38.12	0.37	35.23	0.517
C x S	(25,575)	0.50	1.15	1.17	1.09

(*p < 0.05, **p < 0.01 and ***p < 0.001 after Greenhouse-Geisser correction).

Table 29. Experiment 2b. Results from the RM ANOVA analyses (N=24) for the CA compared to the no-delay CP condition.

	Fixed-interval			Criminal/IS-based	
Effects	df	MSE	F	MSE	F
CONDITION (C)	(1,23)	89.67	4.21*	111.52	4.78*
SITE (S)	(5,115)	12.19	13.58***	16.02	16.47***
PHOTO (P)	(5,115)	38.56	11.75***	44.49	18.34***
C x S	(5,115)	2.89	2.73*	5.90	5.73***
C x P	(5,115)	46.31	11.73***	56.74	12.16***
S x P	(25,575)	0.56	2.81***	1.04	3.86***
C x S x P	(25, 575)	0.67	2.92**	1.12	1.81

(*p < 0.05, **p < 0.01 and ***p < 0.001 after Greenhouse-Geisser correction).

Table 30. Experiment 2b. Results from the RM ANOVA analyses comparing correct identifications in the CP condition to the highest CA lineup member. Note: only subjects that made correct identifications in the CP condition were included in the analyses (Fixed interva,l N=18, Criminal/IS-based N=19).

	Fixed-interval			Criminal/IS-based	
Effects	df	MSE	F	MSE	F
CONDITION	(1,17)	37.41	22.86***	71.80	27.68***
SITE	(5,85)	5.42	10.19***	3.86	8.27***
C x S	(5,85)	4.04	3.65**	1.96	8.37***

(*p < 0.05, **p < 0.01 and ***p < 0.001 after Greenhouse-Geisser correction).

Table 31. Experiment 2b. The lineup member associated with the highest z-score and mean z-score for each participant in the CA condition for the fixed-interval analysis. ID = identification decision. Bolded items represent z-scores and mean z-scores beyond the statistical cut-off. The number of correct rejections are totalled for each ERP-lineup.

Fixed-Interval								
Subject	**Cp3**	**Cpz**	**Cp4**	**P3**	**Pz**	**P4**	**Mean**	**ID**
ERP Line-up A								
04	**IS**	IS	**IS**	**IS**	IS	IS	**IS**	FI
10	F5	F5	F2	F5	F5	F1	F5	CR
14	**F5**	**F5**	**F5**	**F5**	**F5**	**F5**	**F5**	MI
18	F1	F1	F1	F1	F1	F2/F3	F1	CR
21	F5	**F5**	**F5**	F5	**F5**	**F5**	**F5**	MI
28	**F2**	**F2**	**F2**	**F2**	**F2**	**F2**	**F2**	MI
Total$_A$	3	3	2	3	3	3	2	2
ERP Line-up B								
07	F4	**F3**	**F3**	**F3**	**F3**	F3	**F3**	MI
09	**F1**	**F1**	**F1**	**F1**	**F1**	**F1**	**F1**	MI
13	**F3**	**F3**	**F3**	**F3**	**F3**	**F3**	**F3**	MI
25	**F4**	F4	F5	**F4**	F4	F4	F4	CR
26	**F2**	**F4**	**F4**	F2	F4	F4	F4	CR
27	**F3**	**F4**	F5	**F4**	**F4**	F5	F4	CR
Total$_B$	1	1	2	1	2	4	3	3
ERP Line-up C								
02	**F2**	**F2**	**F2**	**F2**	**F2**	**F2**	**F2**	MI
06	**F4**	**F4**	**F4**	**F1**	**F4**	**F1**	**F4**	MI
12	F3	F3	F3	F3	F3	F3	F3	CR
20	**F1**	**F1**	F1	**F1**	F1	F1	**F1**	MI
22	F1	F5	**F5**	F1	F5	**IS**	IS	CR
29	**F3**	**F3**	**F3**	**F3**	**F3**	**F3**	**F3**	MI
Total$_C$	2	2	2	2	3	2	2	2
ERP line-up D								
01	F4	IS	F4	**F4**	F4	F4	F4	CR
05	**F1**	**IS**	**F1**	**F1**	F1	IS	F1/IS	CR
11	**IS**	**IS**	**IS**	**IS**	**IS**	**IS**	**IS**	FI
15	**F4**	F1	F1	**F4**	F4	F1	F4	CR
23	**IS**	**IS**	**IS**	IS	**IS**	**IS**	**IS**	FI
24	**F5**	**F5**	**F5**	**F5**	**F5**	**F5**	**F5**	MI
Total$_D$	1	2	2	1	3	3	3	3
Total$_{ABCD}$	**7**	**8**	**8**	**7**	**11**	**12**	**10**	**10**
%$_{ABCD}$	**29**	**33**	**33**	**29**	**46**	**50**	**42**	**42**
Correct Rejections (CR)								10/24 (42%)
Misidentifications (MI)								11/24 (46%)
False Identifications (FI)								3/24 (13%)

Table 32. Experiment 2b. The lineup member associated with the highest z-score and mean z-score for each participant in the CA condition for the IS-based analysis . ID = identification decision. Bolded items represent z-scores and mean z-scores beyond the statistical cut-off. The number of correct rejections are totalled for each ERP-lineup.

	IS-based							
Subject	**Cp3**	**Cpz**	**Cp4**	**P3**	**Pz**	**P4**	**Mean**	**ID**
ERP Line-up A								
04	**IS**	**IS**	**IS**	**IS**	IS	F2	**IS**	FI
10	**F5**	F5	F2	**F5**	**F5**	F5	F5	CR
14	**F2**	**F5**	**F5**	F2	**F5**	**F5**	**F5**	MI
18	**F1**	F1/F2	F1	**F1**	F2	F1	F1	CR
21	F5	F5	**F5**	F5	F5	**F5**	**F5**	MI
28	**F2**	**F2**	**F2**	**F2**	**F2**	**F2**	**F2**	MI
$Total_A$	1	2	2	2	3	3	2	2
ERP Line-up B								
07	F1	F3	F1	**F5**	IS	**F5**	F5	CR
09	**F1**	**F1**	**F3**	**F1**	**F1**	**F1**	**F1**	MI
13	**F3**	**F3**	**F3**	F3	IS	**F3**	**F3**	MI
25	**F4**	**F5**	**F5**	**F4**	**F4**	F4	F4	CR
26	IS	**F4**	F2	**F2**	F2	F2	F2	CR
27	**F3**	F3	**IS**	**F3**	F3	**IS**	F3	CR
$Total_B$	2	2	2	1	4	2	4	4
ERP Line-up C								
02	**F2**	**F2**	**F2**	**F2**	**F2**	**F2**	**F2**	MI
06	**F4**	**F4**	**F3**	**F1**	**F4**	**F1**	F1	CR
12	IS	F3	IS	IS	F3	F5	F3	CR
20	**F1**	**F1**	F1	**F1**	F1	F2	**F1**	MI
22	F2	IS	**IS**	F2	IS	**IS**	IS	CR
29	**F3**	**F3**	**F3**	**F3**	**F3**	**F3**	**F3**	MI
$Total_C$	2	2	2	2	3	2	3	3
ERP line-up D								
01	F5	F5	**F4**	**F4**	F4	**F4**	F4	CR
05	**IS**	**IS**	**IS**	**IS**	**IS**	**F4**	**IS**	FI
11	**IS**	**IS**	**IS**	**IS**	**IS**	**IS**	**IS**	FI
15	**F2**	**F2**	F2	**F2**	**F2**	**F2**	**F2**	MI
23	**IS**	**IS**	**IS**	**F1**	**F1**	**F1**	IS	CR
24	**F5**	**F5**	**F5**	**F5**	**F5**	**F5**	**F5**	MI
$Total_D$	1	1	1	0	1	0	2	2
$Total_{ABCD}$	**6**	**7**	**7**	**5**	**11**	**7**	**11**	**11**
$\%_{ABCD}$								
Correct Rejections (CR)								11/24 (46%)
Misidentifications (MI)								10/24 (42%)
False Identifications (FI)								3/24 (13%)

Table 33. Experiment 2b. Certainty ratings for the CA condition.

	Certainty Ratings						
Subject	**F1**	**F2**	**F3**	**F4**	**F5**	**IS**	**ID**
ERP Line-up A							
04	0	**7**	0	0	0	0	MI
10	0	0	0	0	**5**	0	MI
14	0	0	0	0	2	0	CR
18	0	1	0	0	2	0	CR
21	0	5	0	0	**8**	0	MI
28	0	**7**	0	0	0	0	MI
M_A	0.0	3.3	0.0	0.0	2.8	0.0	2
ERP Line-up B							
07	**6**	3	5	1	**6**	2	MI
09	0	0	0	0	0	0	CR
13	0	0	0	0	0	0	CR
25	4	0	1	0	1	1	CR
26	**5**	0	**5**	0	0	0	MI
27	4	3	0	0	1	0	CR
M_B	3.2	1.0	1.8	0.2	1.3	0.5	4
ERP Line-up C							
02	5	**7**	0	0	2	4	MI
06	0	0	0	0	3	0	CR
12	0	1	**5**	1	0	0	MI
20	2	1	2	3	0	2	CR
22	**5**	0	0	0	0	0	MI
29	0	0	**6**	3	0	0	MI
M_C	2.0	1.5	2.2	1.2	0.8	1.0	2
ERP Line-up D							
01	0	0	0	0	**5**	2	MI
05	0	0	1	0	0	0	CR
11	0	0	0	0	0	3	CR
15	1	0	1	4	1	1	CR
23	1	1	0	0	0	0	CR
24	0	0	0	0	**5**	0	MI
M_D	0.3	0.2	0.3	0.7	1.8	1.0	4
$\boldsymbol{M_{ABCD}}$	**1.4**	**1.2**	**1.1**	**0.5**	**1.7**	**0.8**	**12**
Correct Rejections (CR)							12/24 (50%)
Misidentifications (MI)							12/24 (50%)
False Identifications (FI)							0/24 (0%)

Table 34. Experiment 2b. Percentage of times participants classified each lineup member as the criminal in the criminal-absent condition. * denotes when participants refrained from making a button press response more than 10% of the time. ID = identification decision.

Button Press							
Subject	**F1**	**F2**	**F3**	**F4**	**F5**	**IS**	**ID**
ERP Line-up A							
04	0	0	0	0	0	0*	CR
10	0	0	0	0	0	0	CR
14	0	0	0	0	0	0	CR
18	0	0	0	0	0	0	CR
21	0	0 *	0	0	**77.5***	5*	MI
28	2.5	10*	0	0	0*	0*	CR
M_A	0.4	1.7	0.0	0.0	12.9	0.8	5
ERP Line-up B							
07	0	0	0	0	0	0	CR
09	0	0	0	0	0	0	CR
13	0	0	0	0	0	0	CR
25	0	0	0	0	0	0	CR
26	0	0	0	0	0	0	CR
27	0	0	0	0	0	0	CR
M_B	0	0	0	0	0	0	6
ERP Line-up C							
02	5	**92.5**	2.5	0	0	0	MI
06	0	0	0	7.5*	0	0	CR
12	0	0*	7.5*	0*	0	0	CR
20	0	0	7.5	2.5	0	0	CR
22	0	0	0	0	0	0	CR
29	0	0	**80**	0	0	0	MI
M_C	0.8	15.4	15.0	0.4	0	0	4
ERP Line-up D							
01	0	0	0	0	5	10	CR
05	0	0	0	0	0	0	CR
11	0	0	0	0	0	0	CR
15	0	0	0	0	0	0	CR
23	0	0	0	0	0	0	CR
24	0	0	0	0	**97.5**	0	MI
M_D	0	0	0	0	17.1	1.7	5
M_{ABCD}	0.3	4.3	3.8	0.1	7.5	0.6	20
Correct Rejections (CR)							20/24 (83%)
Misidentifications (MI)							4/24 (17%)
False Identifications (FI)							0/24 (0%)

Table 35. Experiment 2b. Summary chart of identification decisions for the CA condition.

Identification Decisions			
Subject #	**Certainty Ratings**	**Button Press**	**ERP Patterns**
ERP Line-up A			
04	MI	CR	FI
10	MI	CR	CR
14	CR	CR	MI
18	CR	CR	CR
21	MI	MI	MI
28	MI	CR	MI
$Total_A$	2	5	2
ERP Line-up B			
07	MI	CR	CR
09	CR	CR	MI
13	CR	CR	MI
25	CR	CR	CR
26	MI	CR	CR
27	CR	CR	CR
$Total_B$	4	6	4
ERP Line-up C			
02	MI	MI	MI
06	CR	CR	CR
12	MI	CR	CR
20	CR	CR	MI
22	MI	CR	CR
29	MI	MI	MI
$Total_C$	2	4	3
ERP line-up D			
01	MI	CR	CR
05	CR	CR	FI
11	CR	CR	FI
15	CR	CR	MI
23	CR	CR	CR
24	MI	MI	MI
$Total_D$	4	5	2
Correct Rejections (CR)	12/24 (50%)	20/24 (83%)	11/24 (46%)
Misidentifications (MI)	12/24 (50%)	4/24 (17%)	10/24 (42%)
False Identifications (FI)	0/24 (0%)	0/24 (0%)	3/24 (13%)

Appendix B: Figures

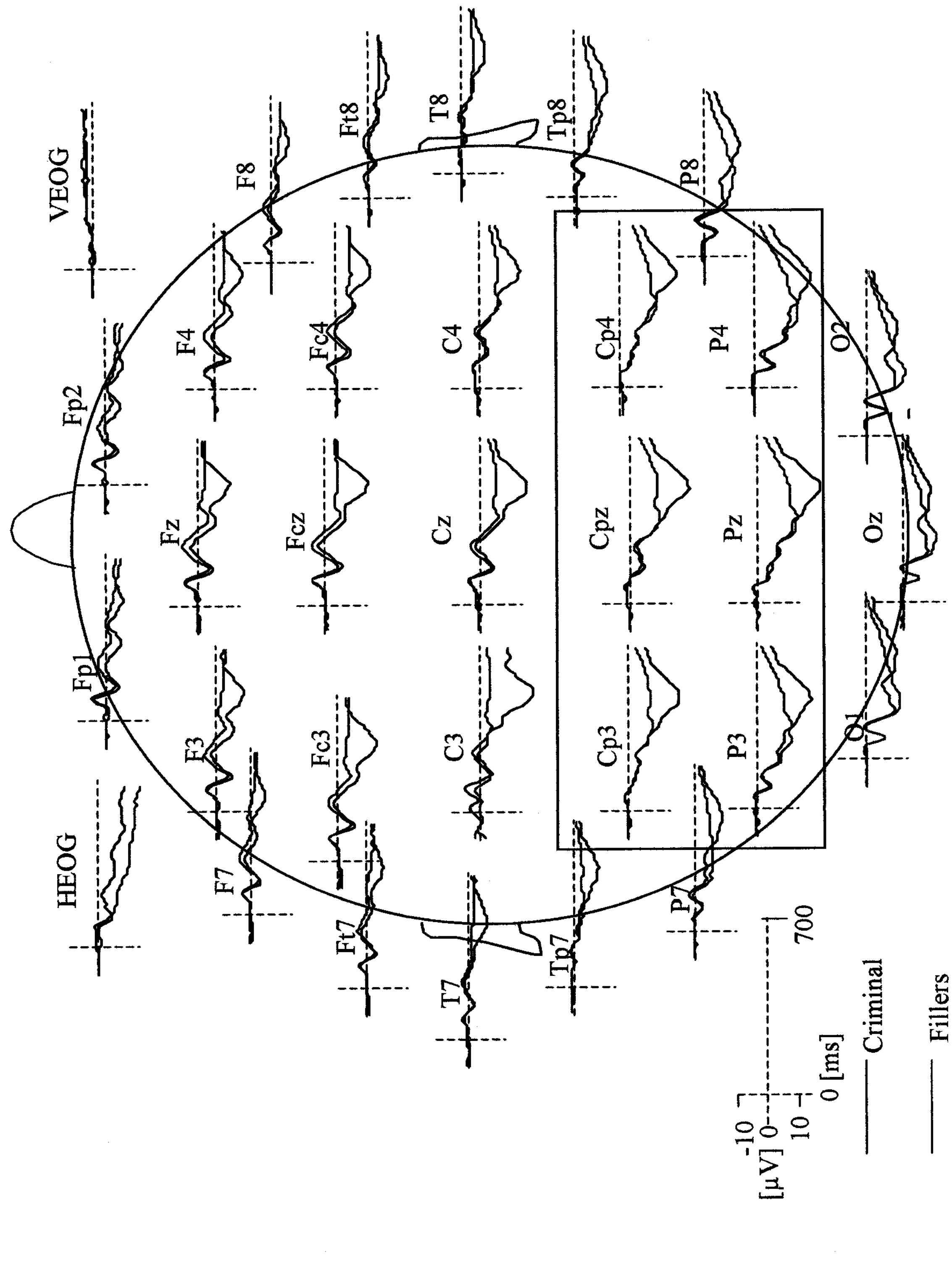

Figure 1 Grand average waveforms across the 30 scalp sites for the standard condition.

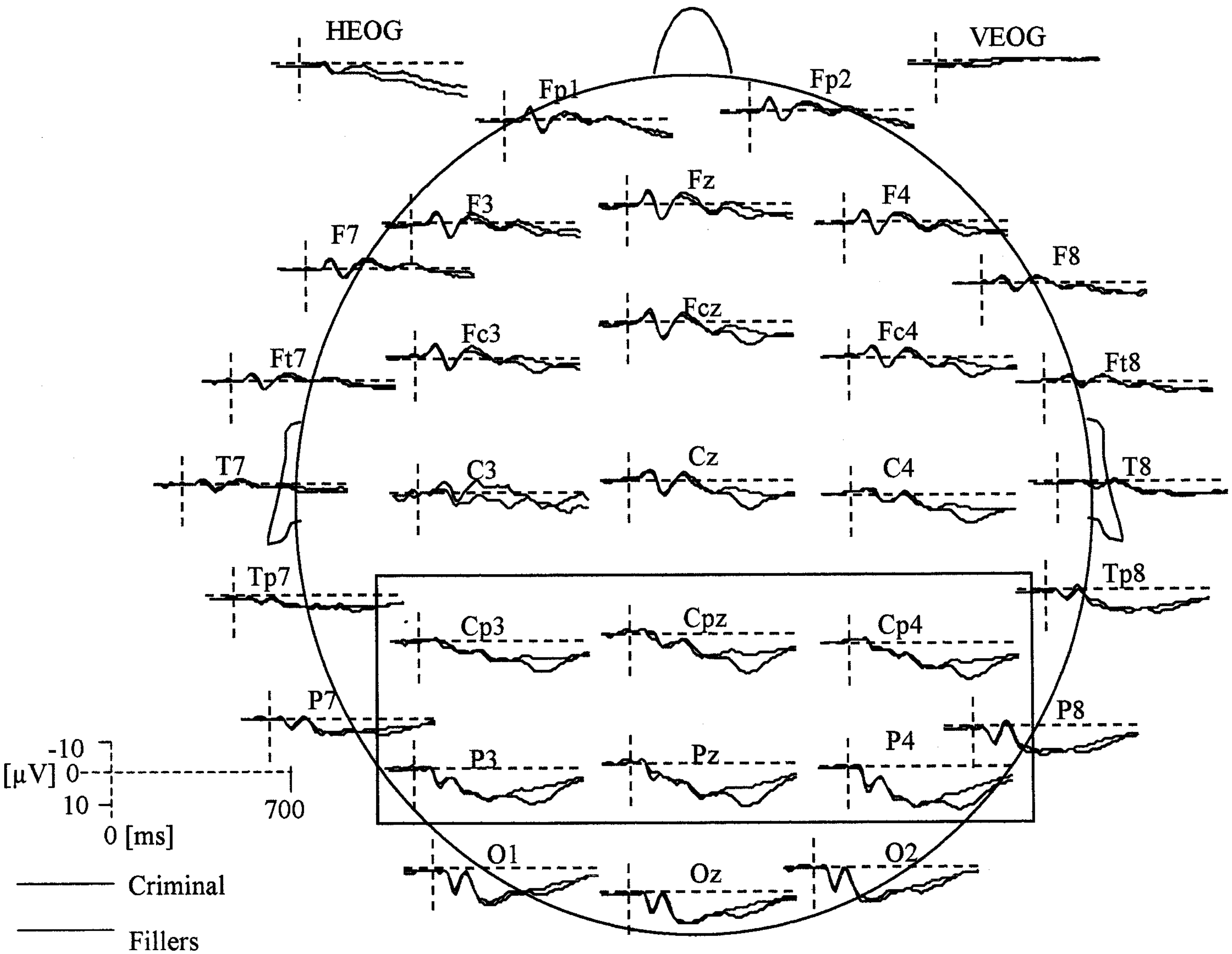

Figure 2 Grand average waveforms across the 30 scalp sites for the deception condition.

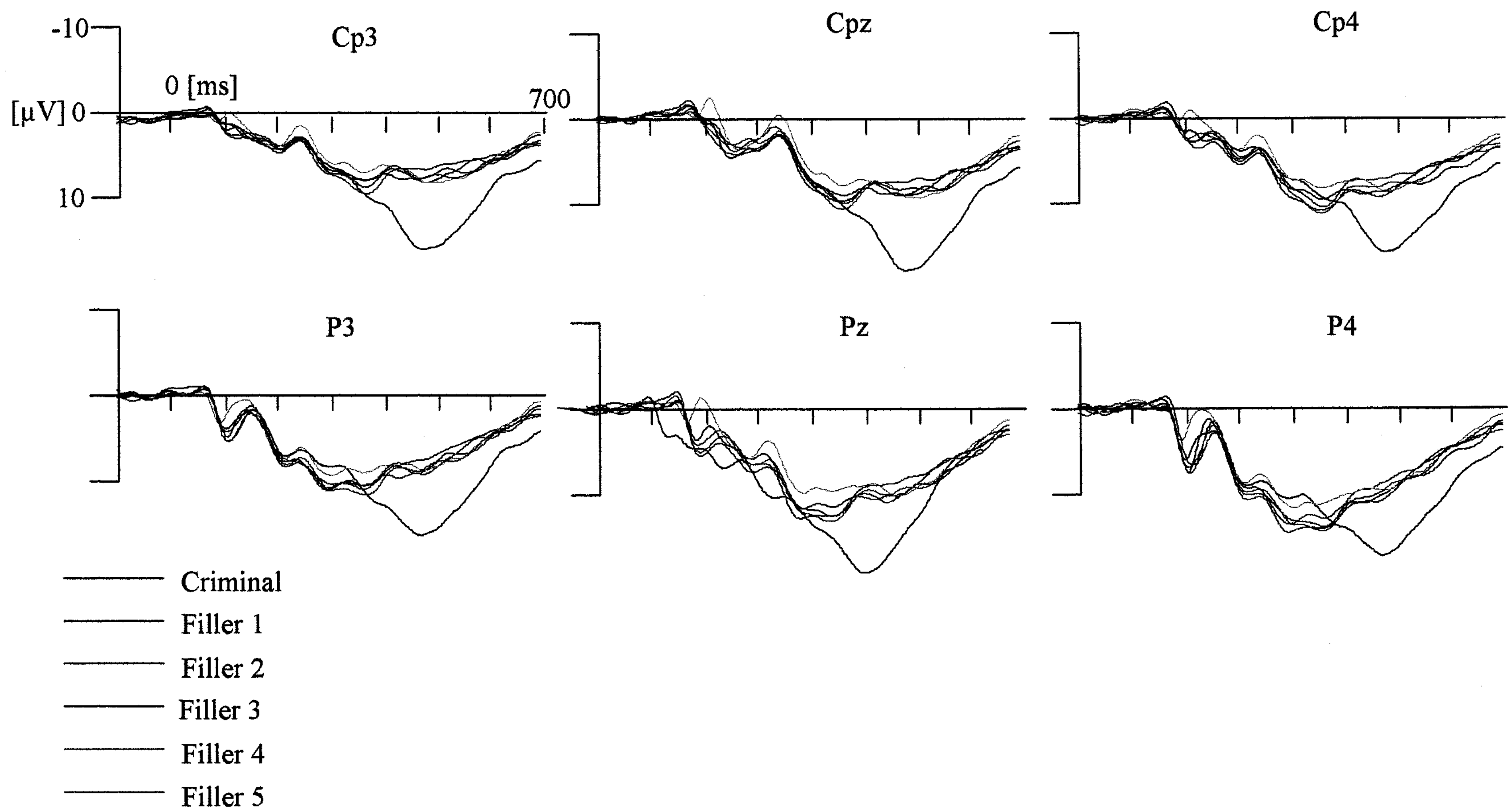

Figure 3 Grand average waveforms across the 6 centro-parietal electrode sites for each lineup member in the standard condition.

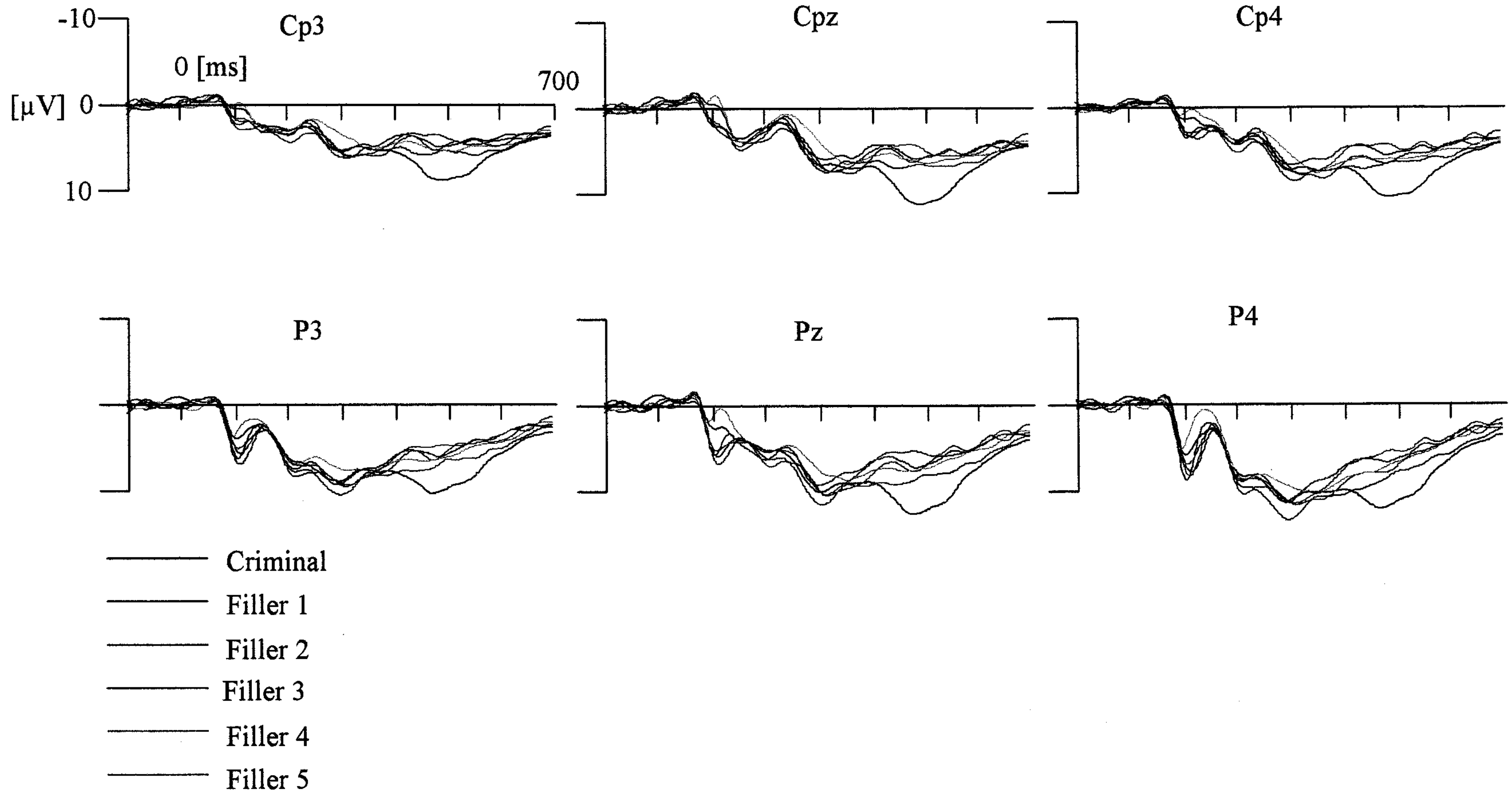

Figure 4 Grand average waveforms across the 6 centro-parietal electrode sites for each lineup member in the deception condition.

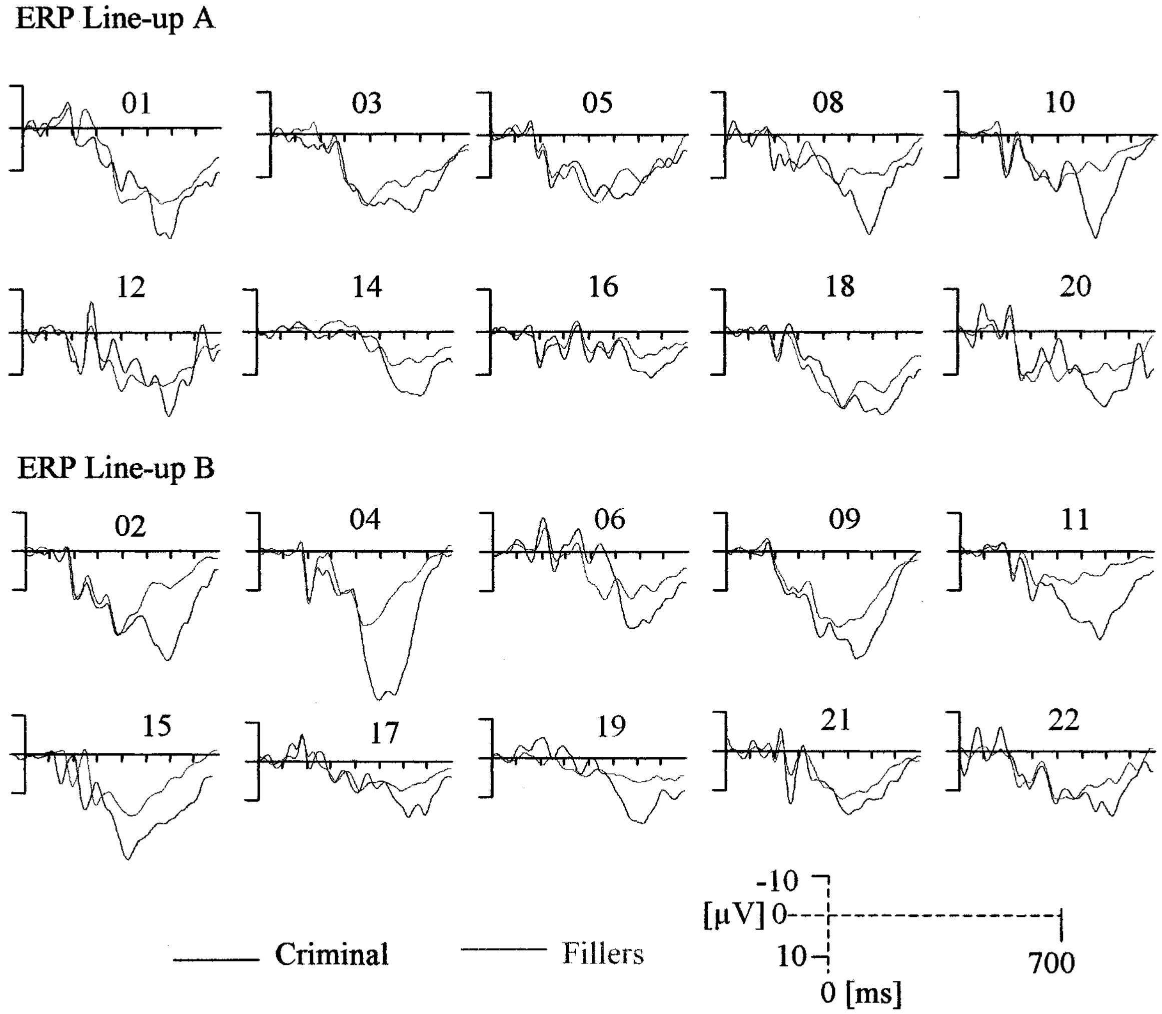

Figure 5 Average waveforms for each participant at the Pz electrode for the standard condition.

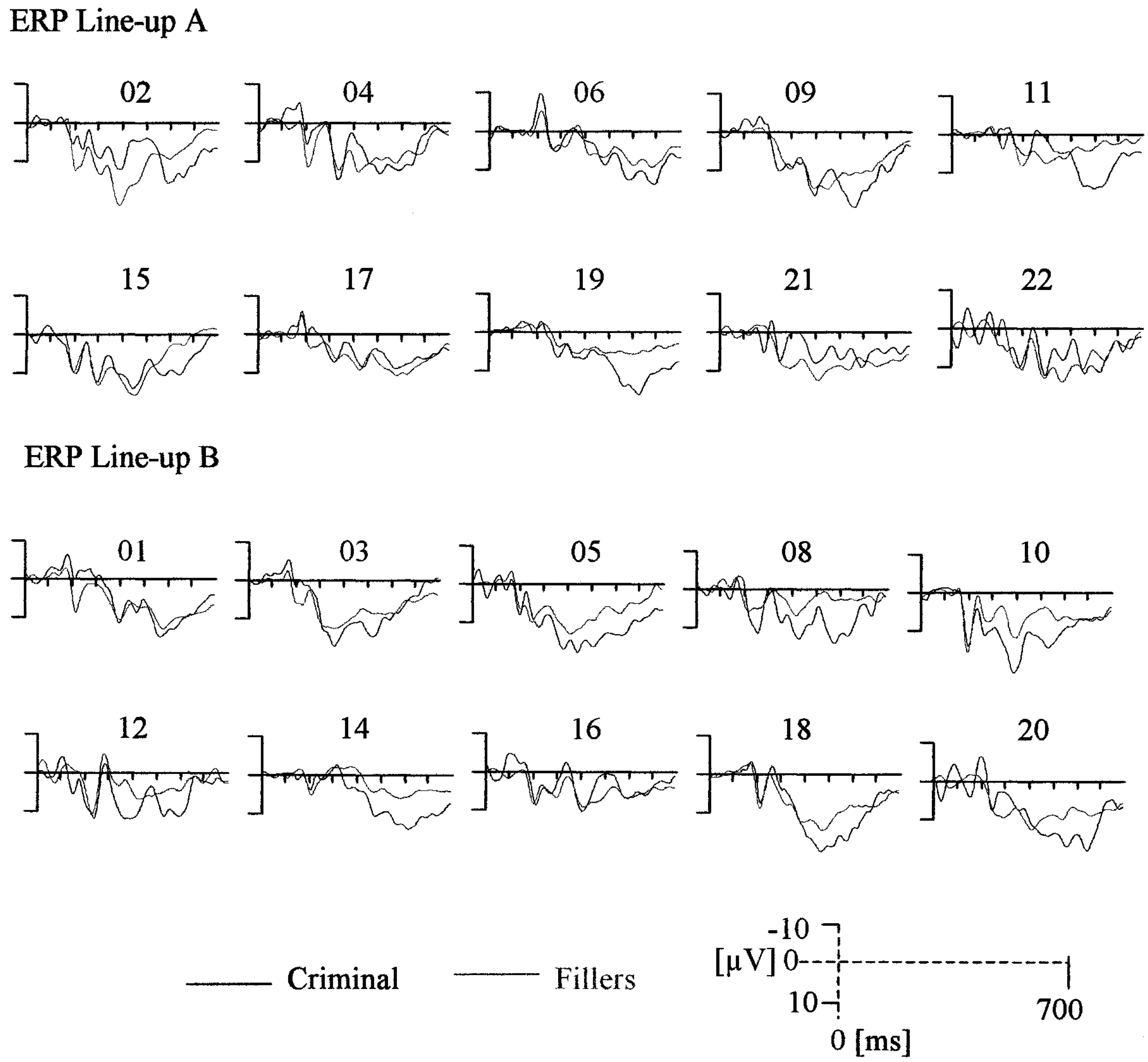

Figure 6. Average waveforms for each participant at the Pz electrode for the deception condition.

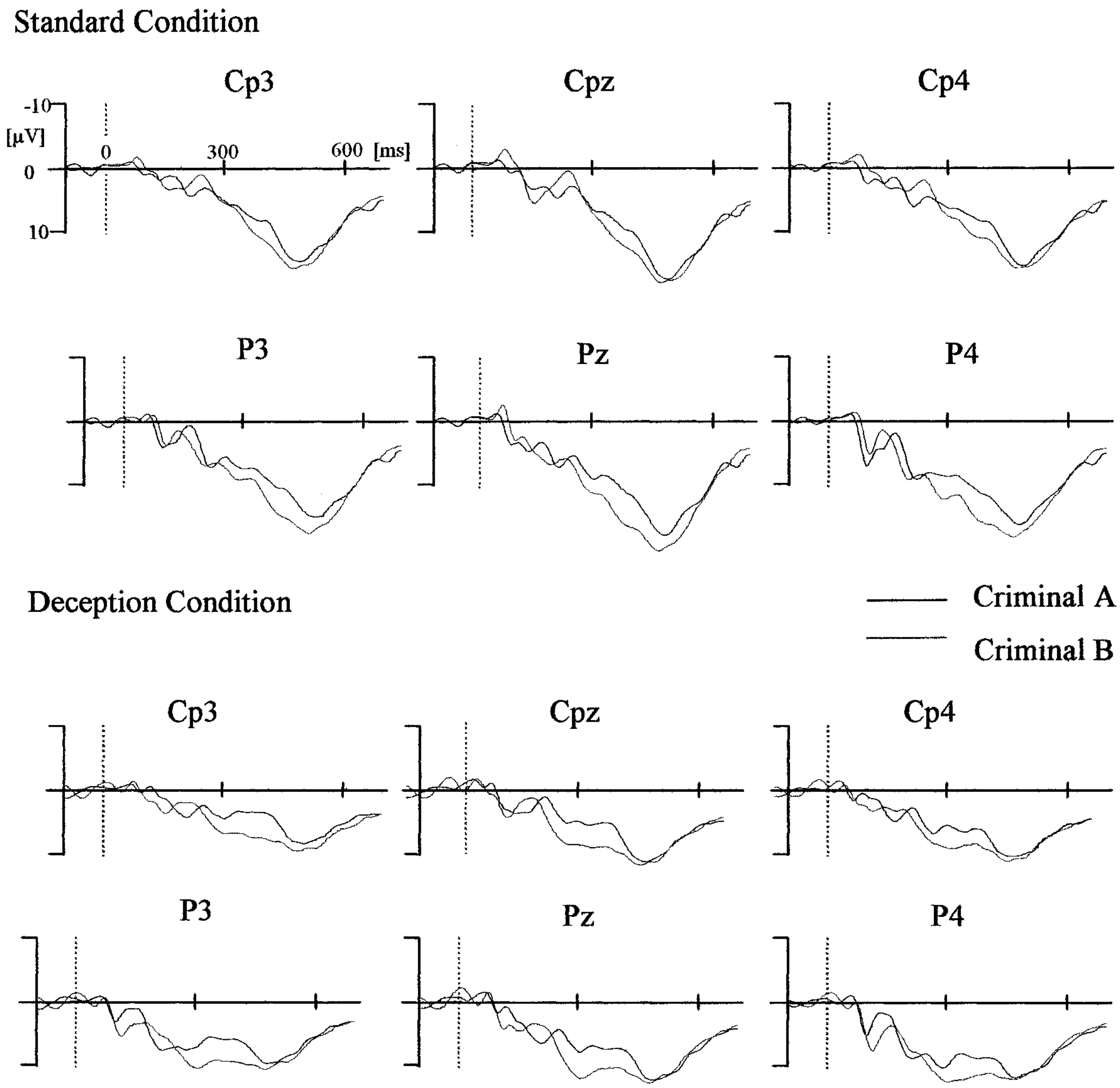

Figure 7 Grand average waveforms for criminal A and B for the standard and deception conditions.

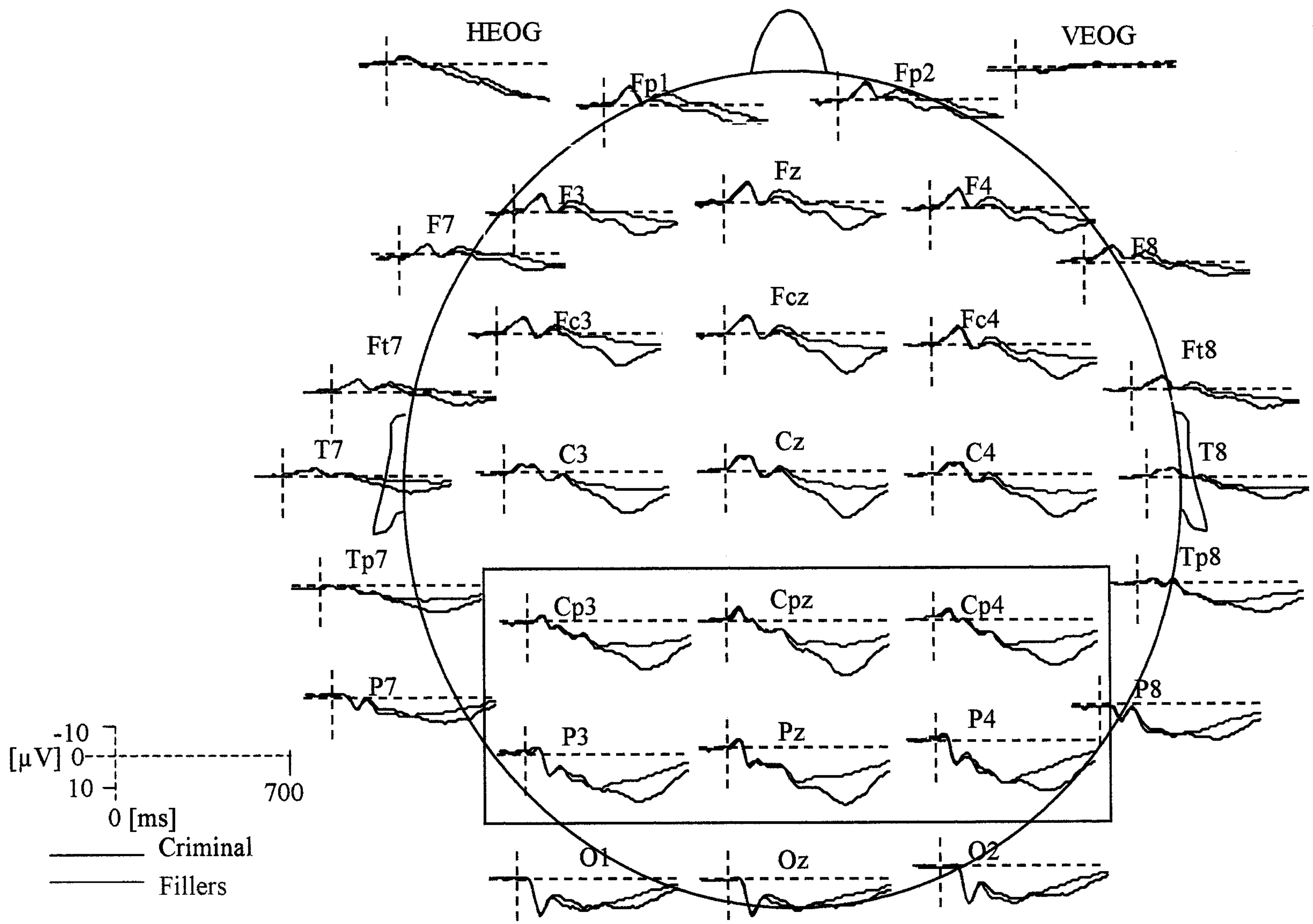

Figure 8 Grand average waveforms across the 30 scalp sites for the no-delay condition.

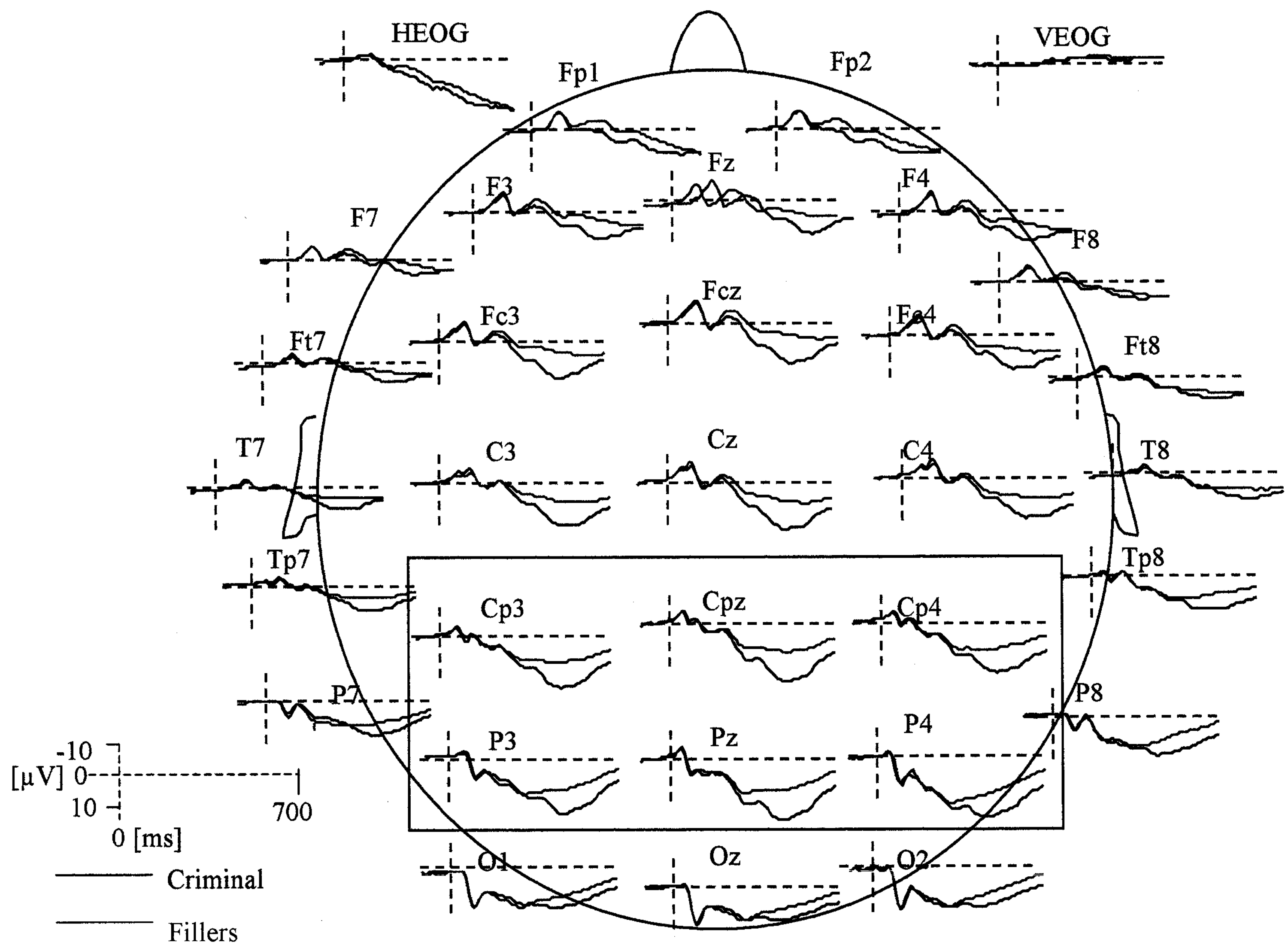

Figure 9 Grand average waveforms across the 30 scalp sites for the 1-hour delay condition.

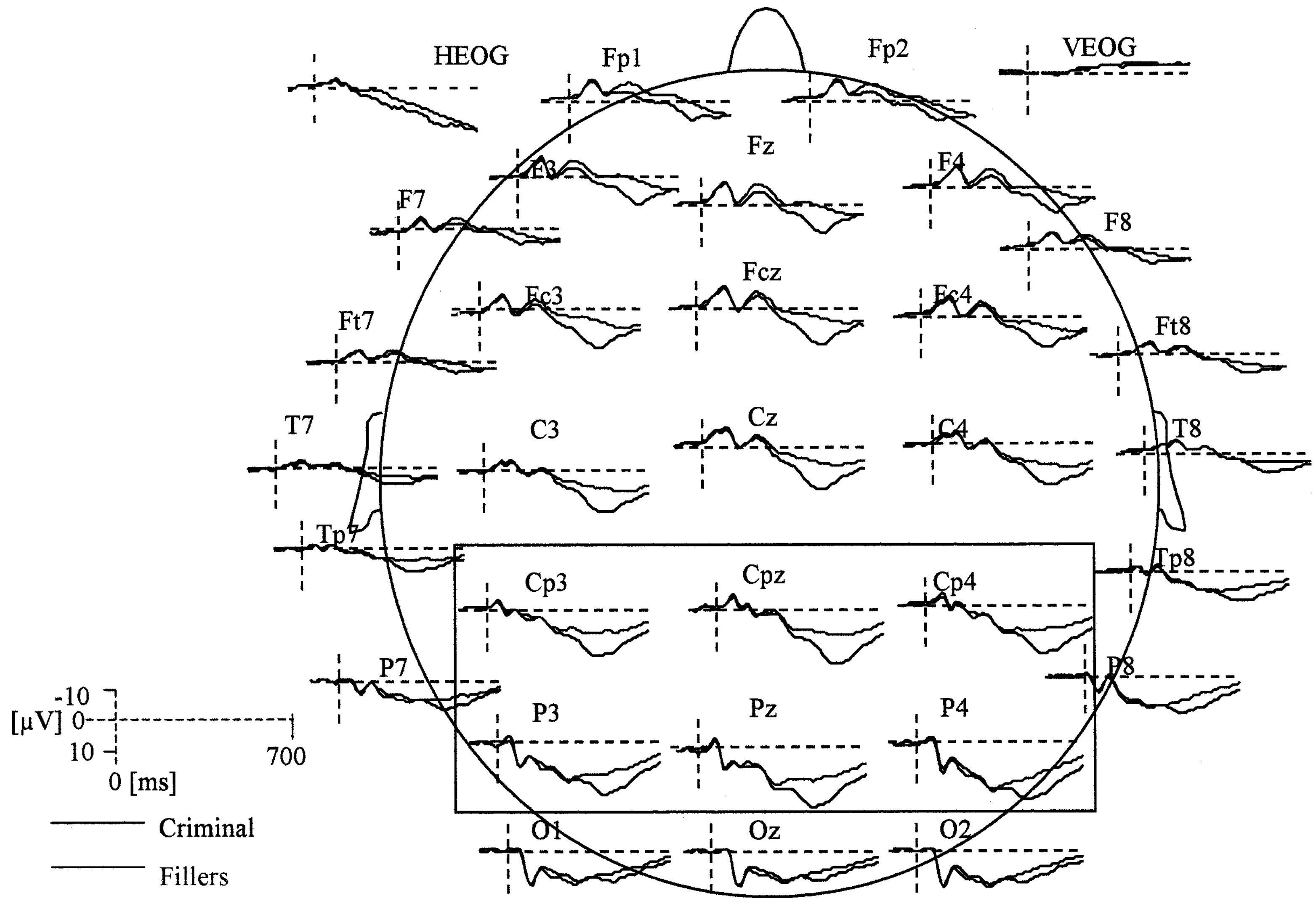

Figure 10 Grand average waveforms across the 30 scalp sites for the 1-week delay condition.

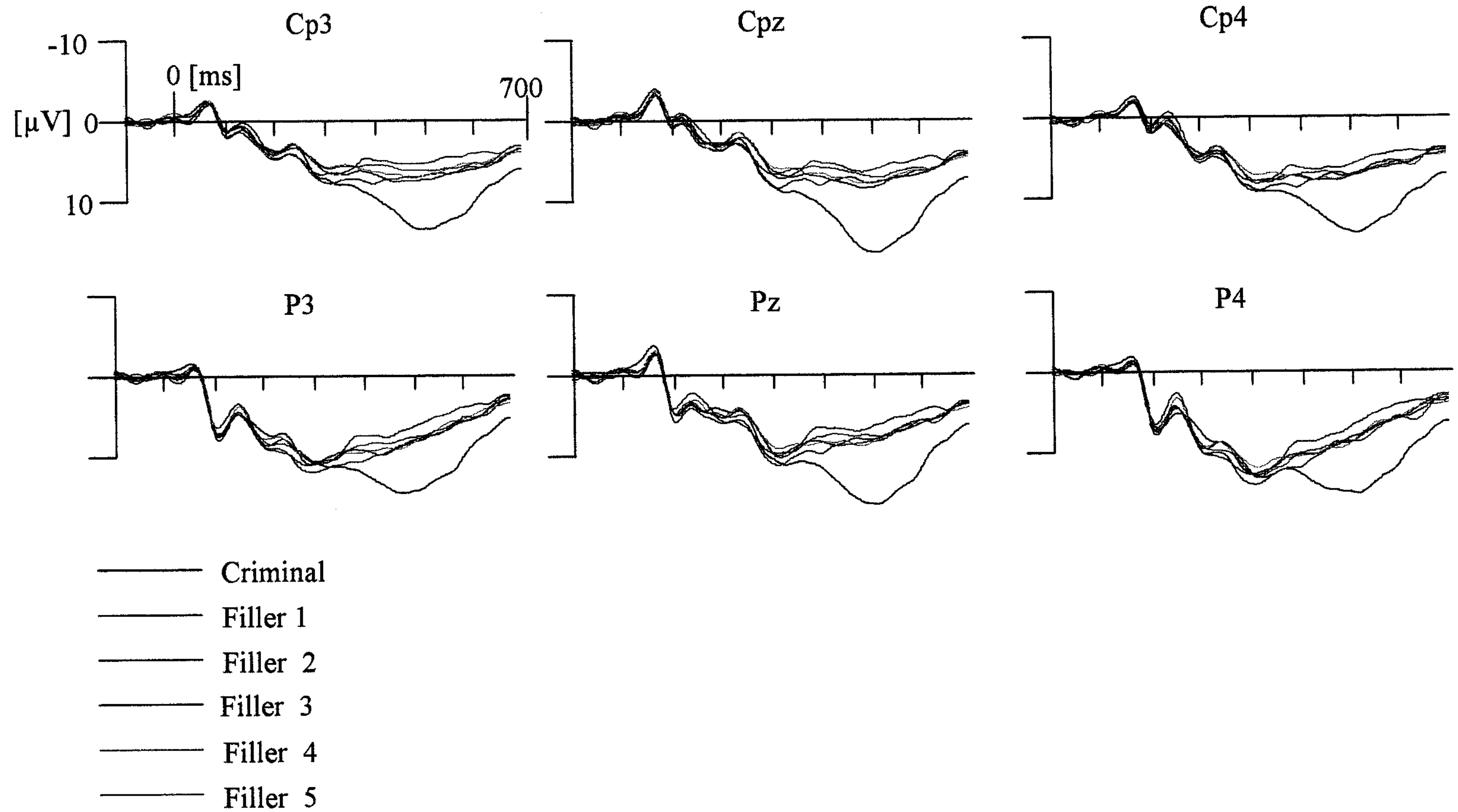

Figure 11 Grand average waveforms for each lineup member in the no-delay condition.

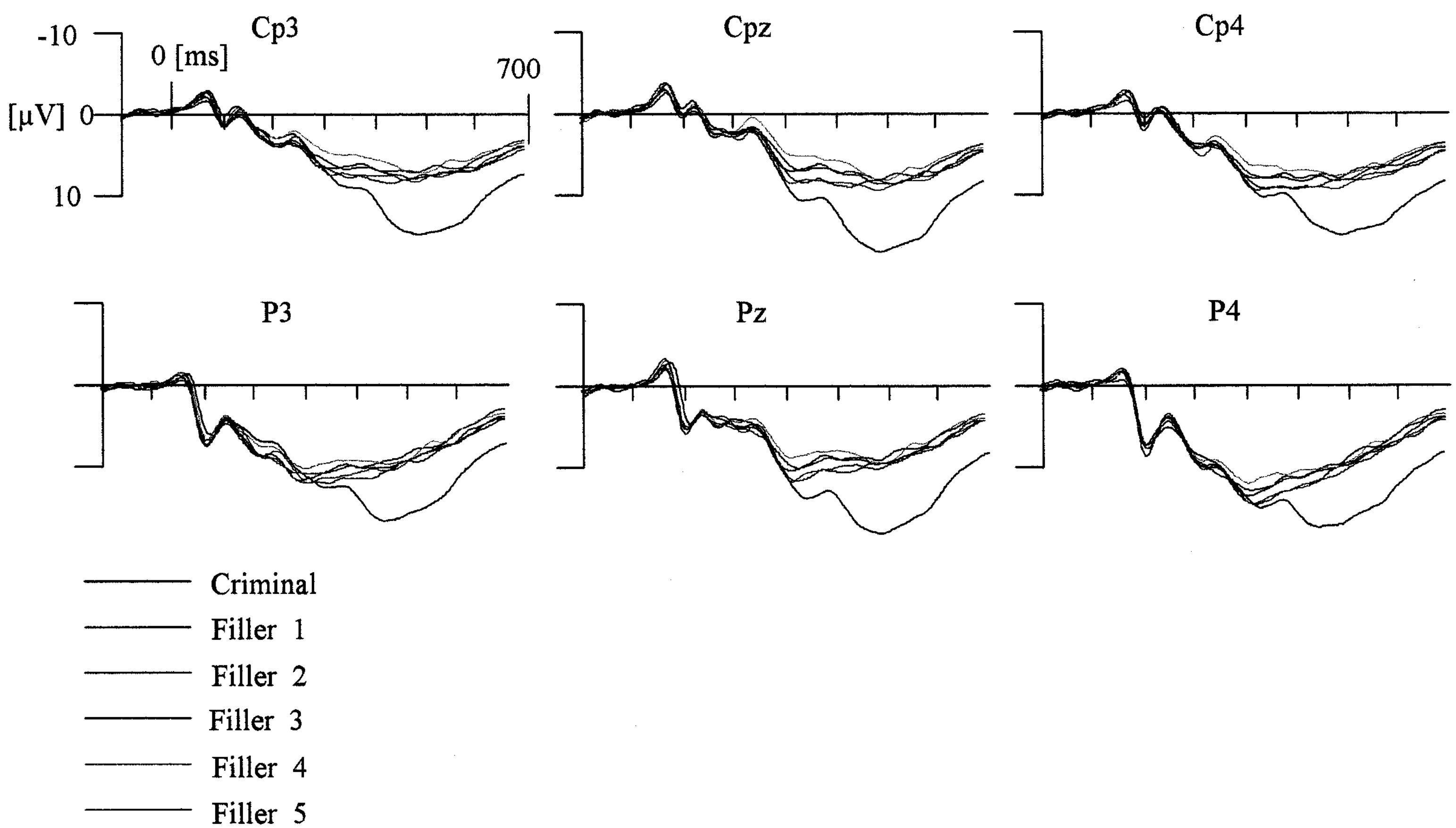

Figure 12 Grand average waveforms for each lineup member in the 1-hour delay condition.

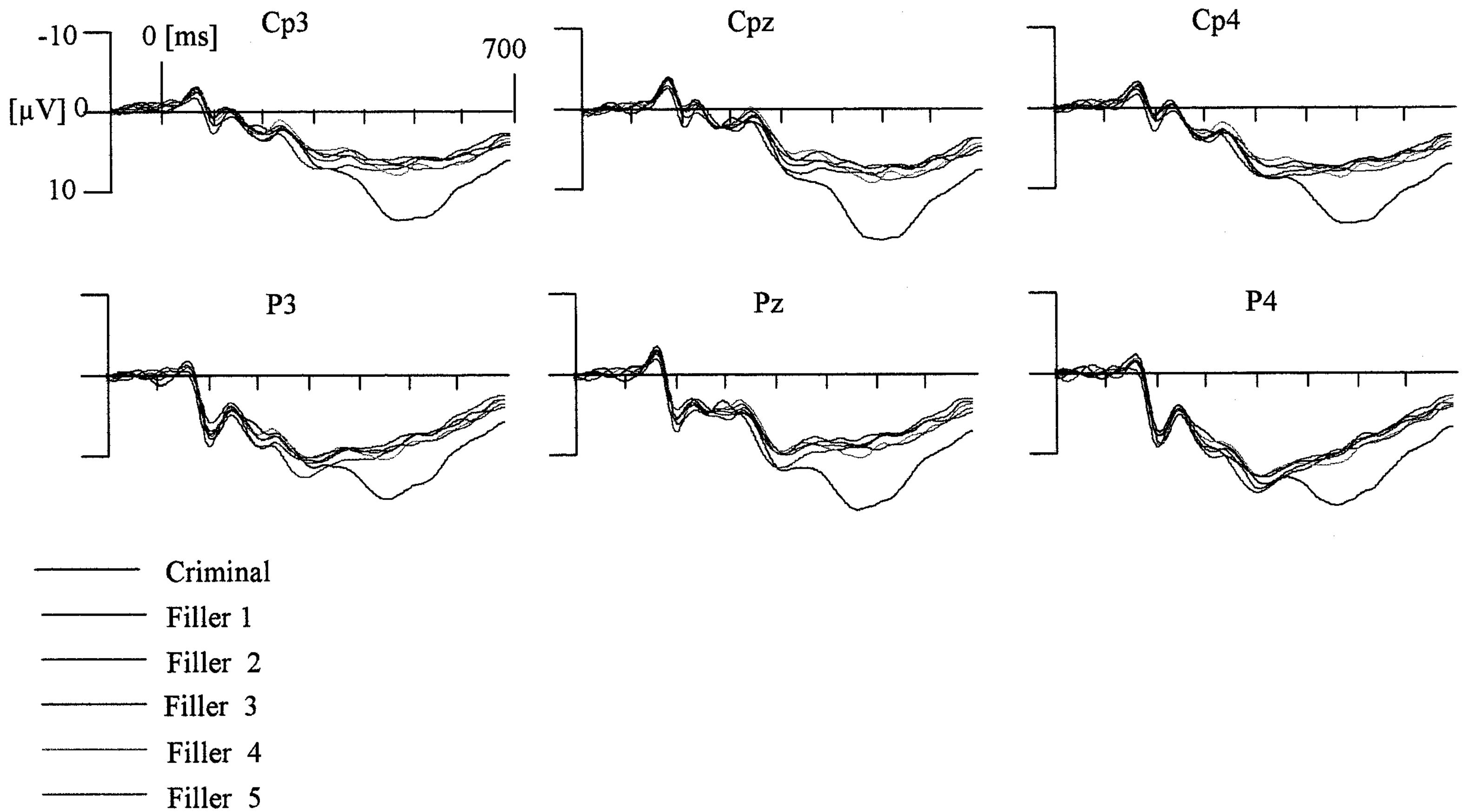

Figure 13 Grand average waveforms for each lineup member in the 1-week delay condition.

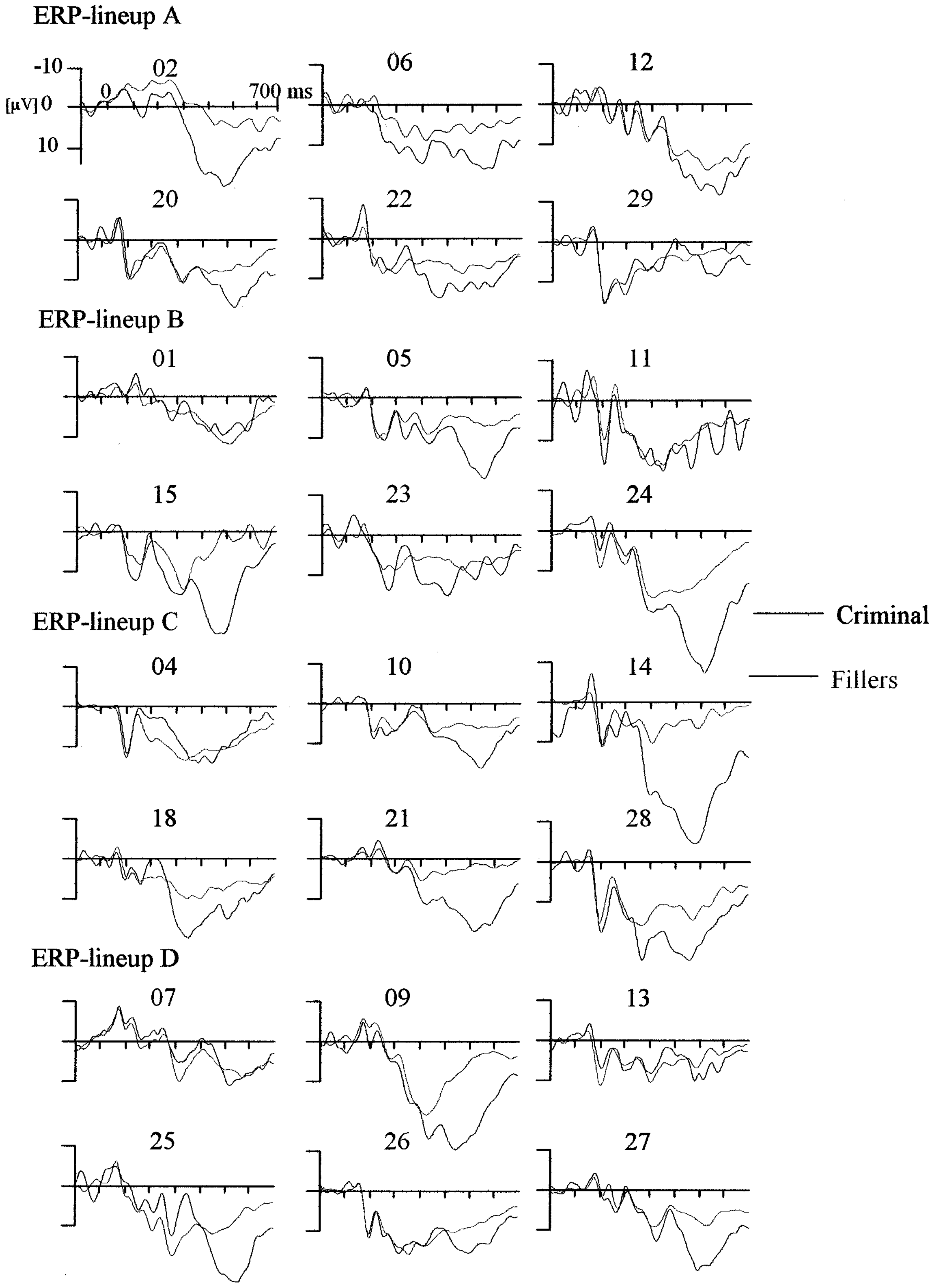

Figure 14 Average waveforms for each participant at the Pz electrode for the no-delay condition.

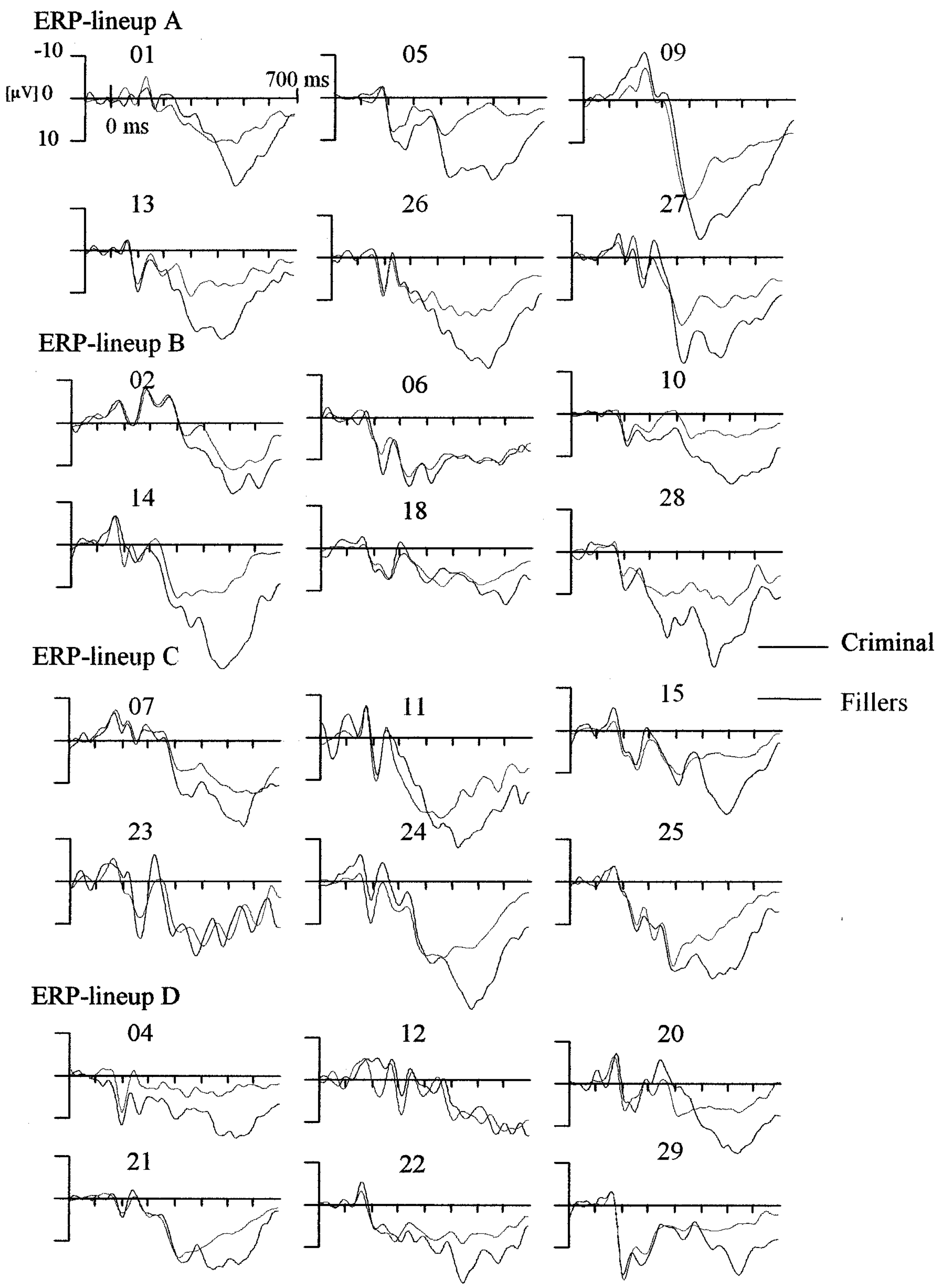

Figure 15 Average waveforms for each participant at the Pz electrode for the 1-hour delay condition.

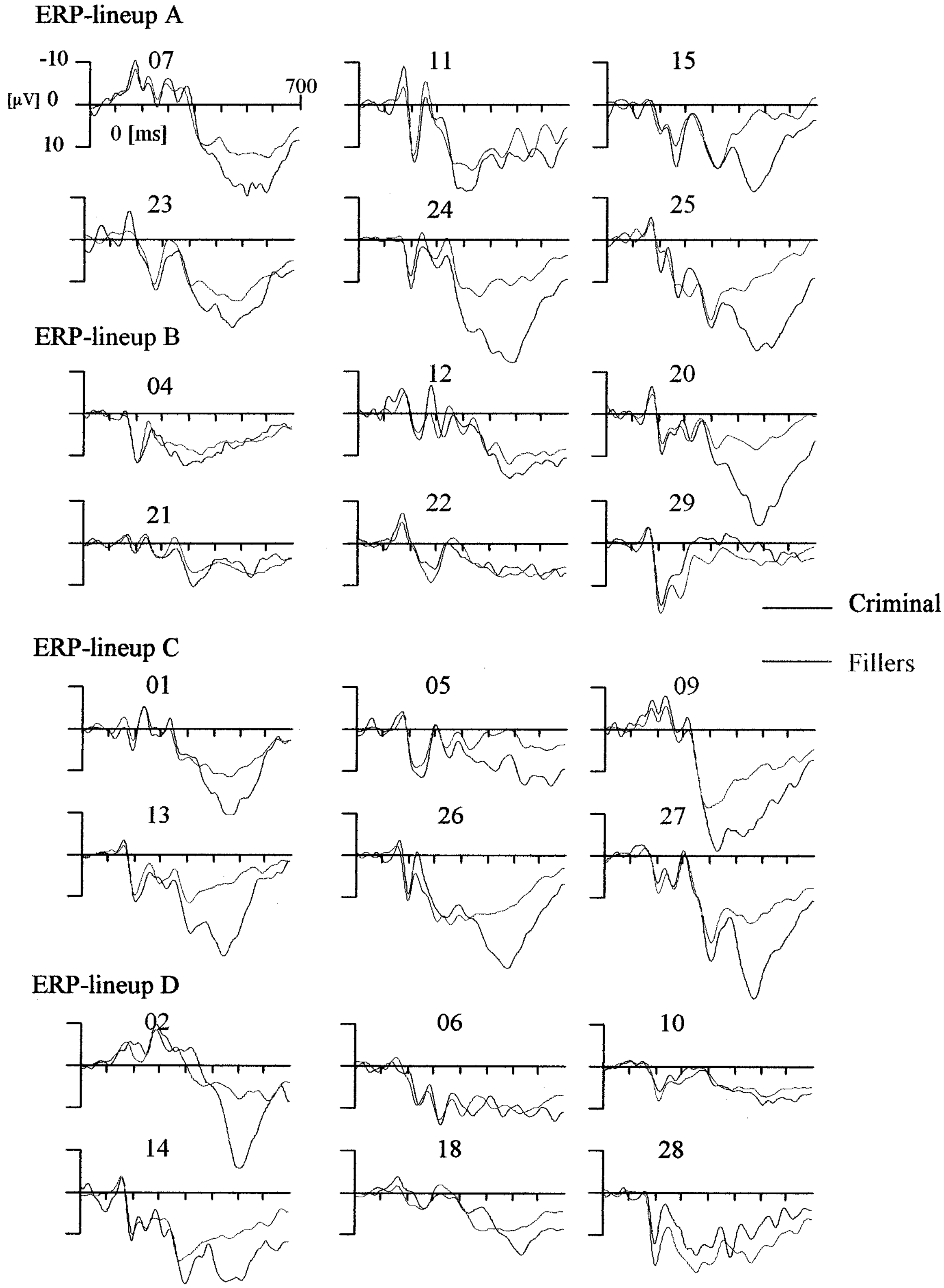

Figure 16 Average waveforms for each participant at the Pz electrode for the 1-week delay condition.

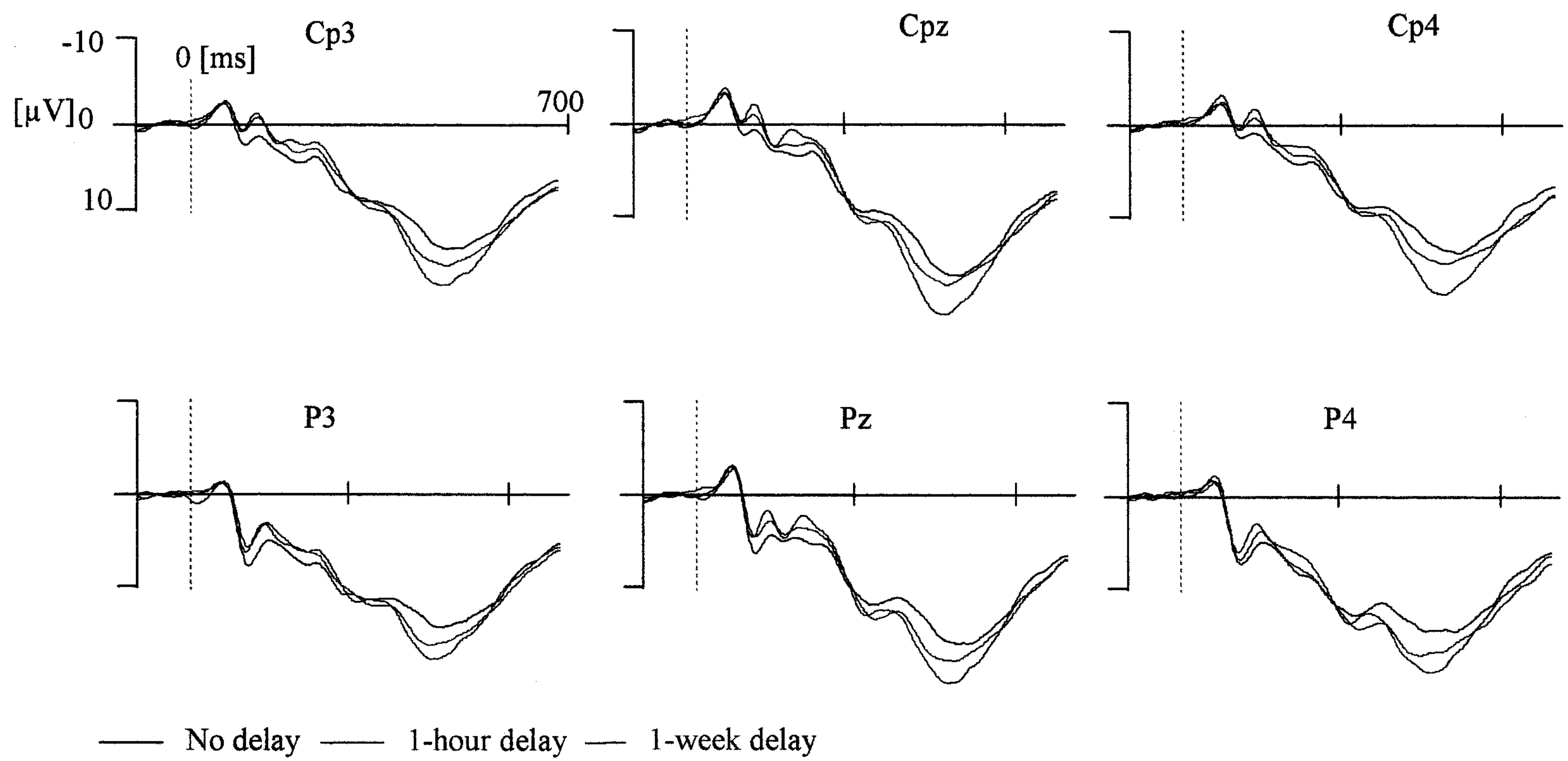

Figure 17 Grand average waveforms to the criminal for participants that made correct identifications according to the criminal-based analyses, (note: no-delay n = 19, 1-hour delay n = 20 , 1-week delay n = 14).

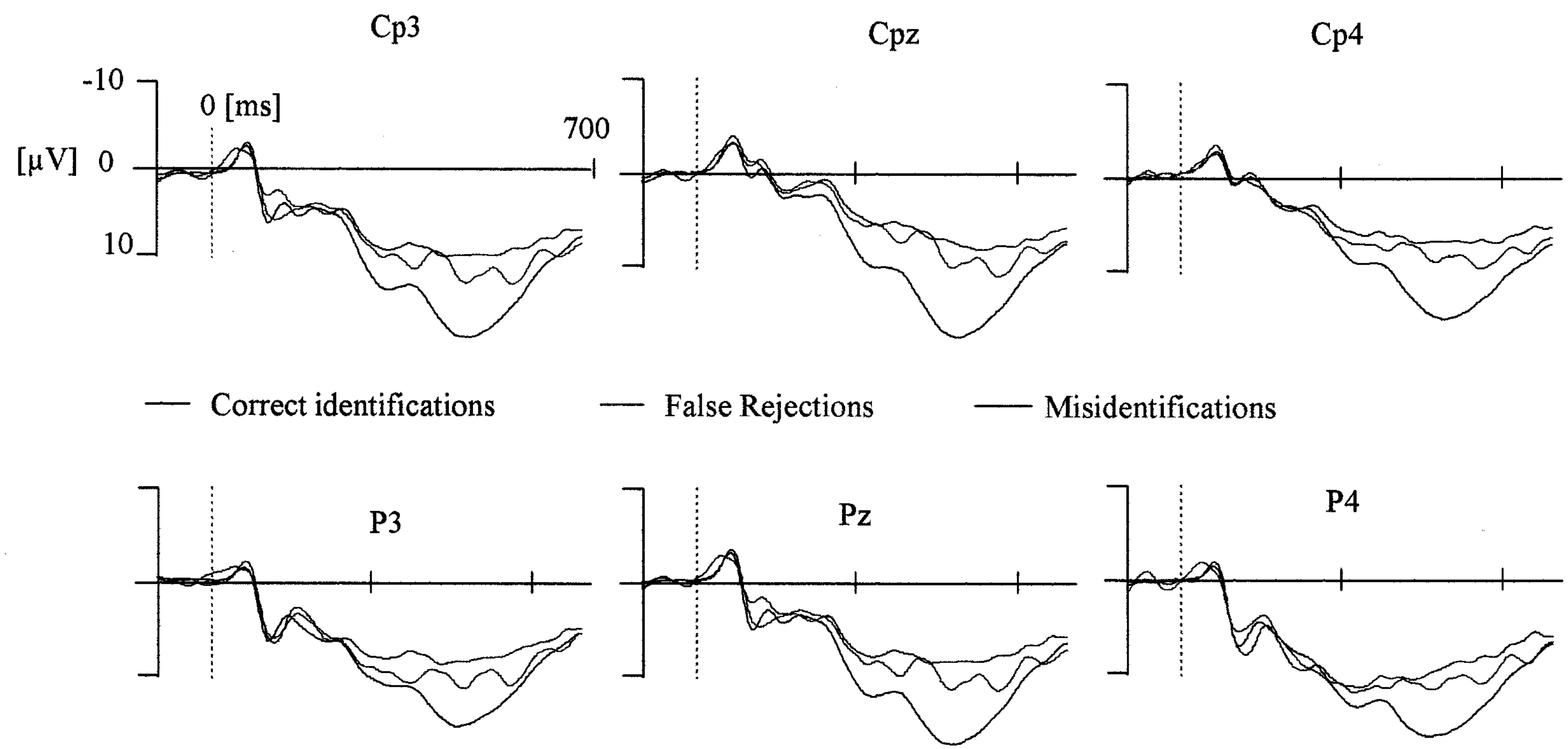

Figure 18 Grand average waveforms elicited to the criminal when participants made either 1) correct identifications (CI), 2) false rejections (FR) or 3) misidentifications (MI), irrespective of time delay. Note for CI: n = 53, FR: n = 8, MI: n = 11.

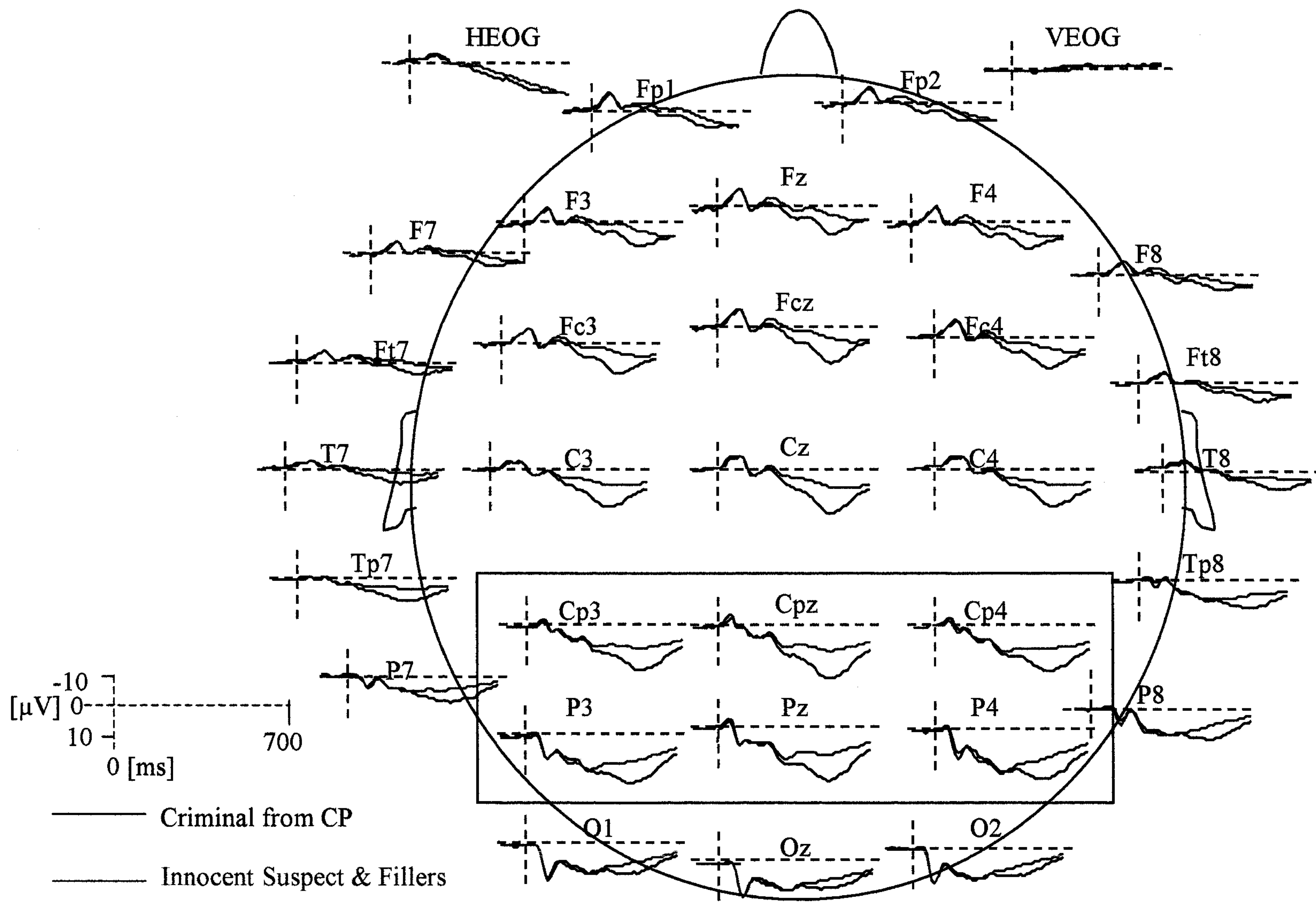

Figure 19 Grand average waveforms across the 30 scalp sites for the lineup members in the CA condition compared to the criminal from the CP conditions.

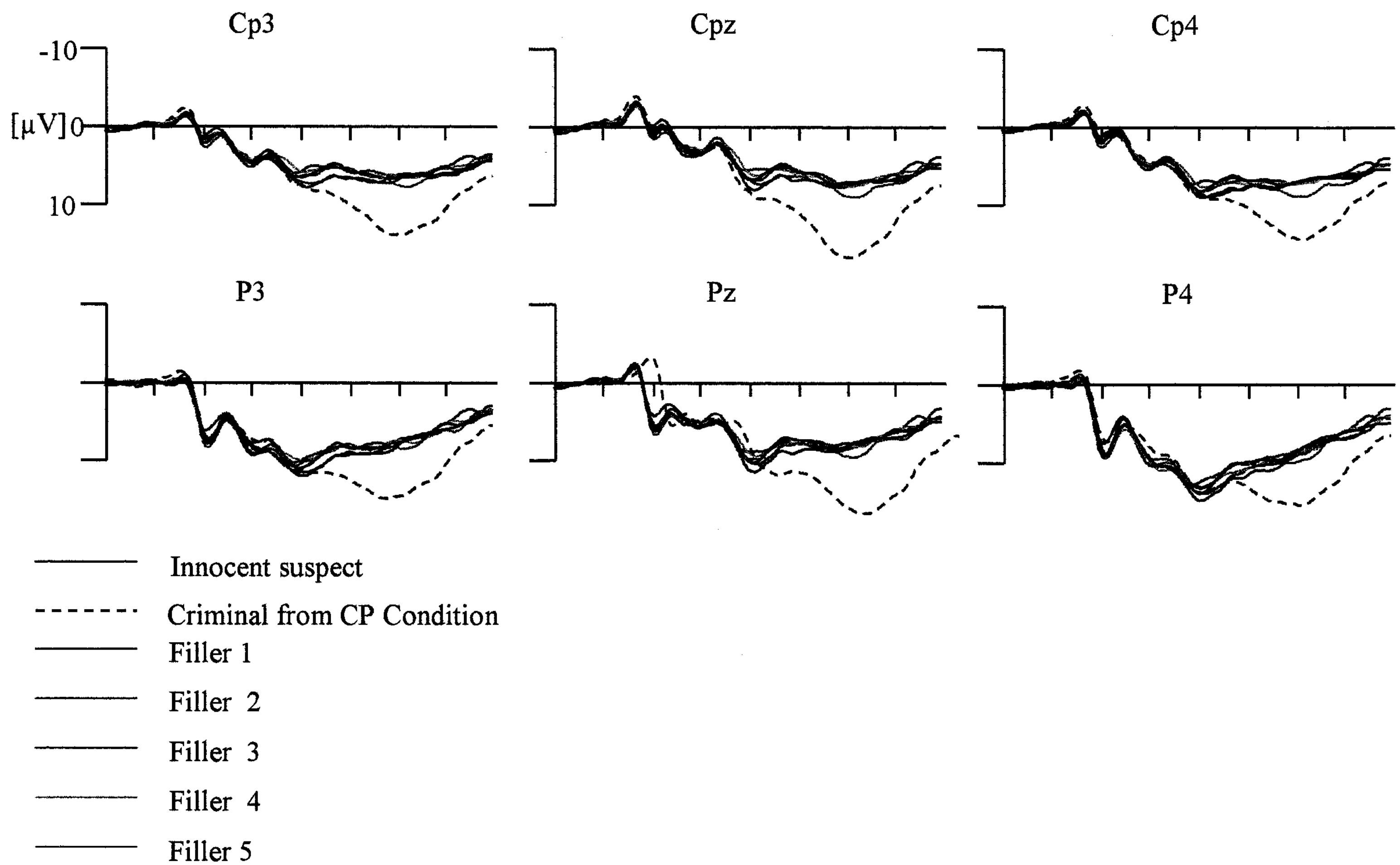

Figure 20 Grand average waveforms for the CA condition for each lineup member compared to the criminal in the CP condition.

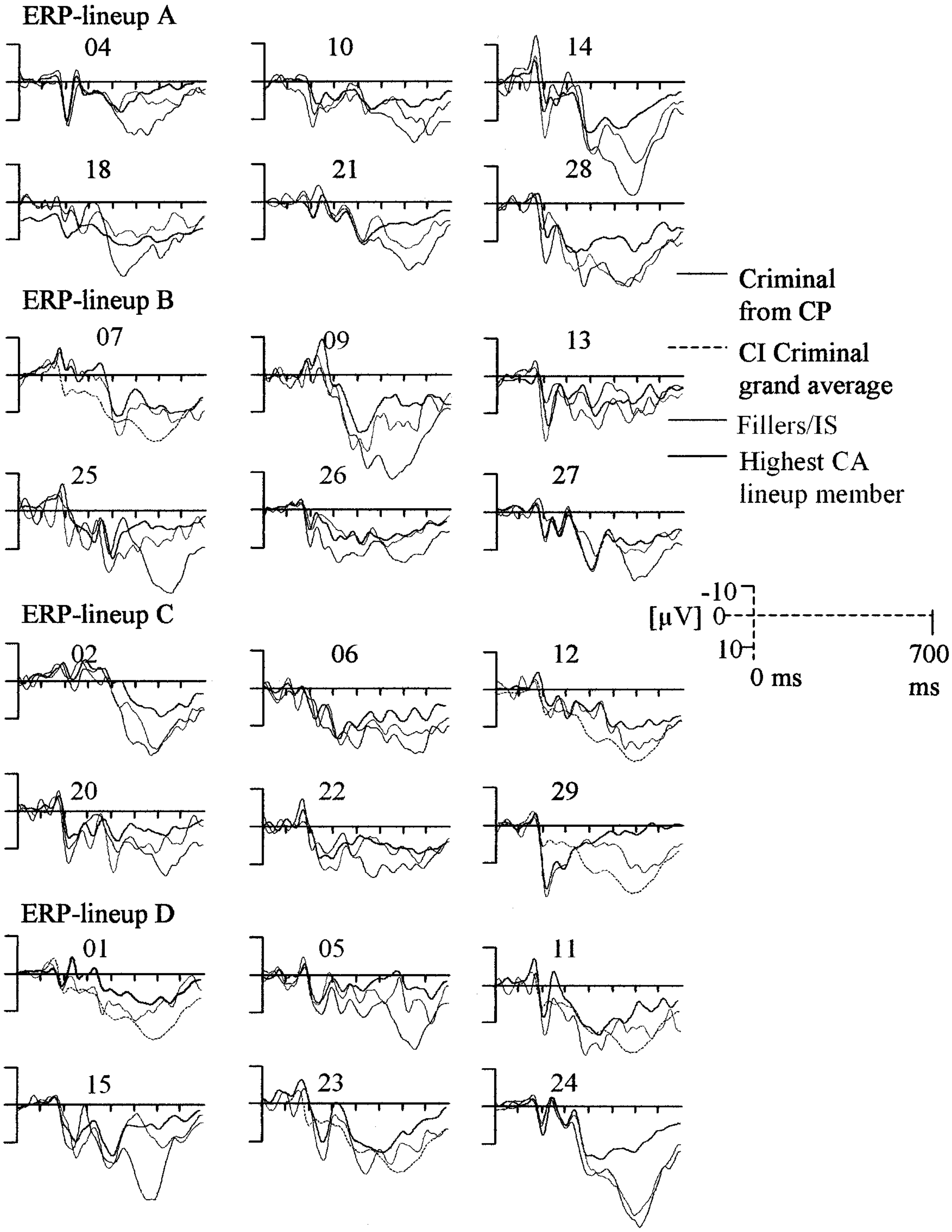

Figure 21 Average waveforms for each participant at the Pz electrode for the highest CA member compared to the remaining fillers combined and to either their waveform to the criminal from the CP condition (if participants made a CI) or to the grand average criminal waveform from the CP condition.

Appendix C: Lineup member selection and assessment of lineup bias using a mock witness procedure

In 1999, a guidebook was published outlining national best practice guidelines for the administration of live and photograph lineups (Reno, Fisher, Robinson, Brennan, & Travis, 1999). This guide involves four major Rules and has subsequently been updated in 2003 (Ashcroft, Daniels, & Hart, 2003). One of the Rules states that the additional members of the lineup (referred to as fillers) should be selected based on overlapping features with the criminal or from eyewitnesses' verbal description of the criminal in terms of approximate height, weight, hair features, age and gender.

Following lineup filler selection, a mock witness procedure (Doob & Kirshenbaum, 1973) has been developed for the assessment of lineup bias (Wells, 1998). For this procedure, without ever actually seeing the crime, *mock witnesses* are provided with written descriptions of the criminal and are asked to try and identify the criminal from the lineup. A lineup is considered biased if the criminal can reliably be identified over the other lineup members. Before the initiation of the Experiment 2a and 2b, it was deemed important to follow the recommended eyewitness procedure guidelines for lineup member selection and to ensure that the lineups were not biased based on the mock witness procedure.

Methods

Lineup member selection

Participants, stimuli and procedures

Five participants watched four 60 second non-violent simulated crime videos (described in the Methods section for Experiment 2a), one for each of the four corresponding ERP-lineups. Following each video, the five participants were asked to write an open-ended paragraph describing the criminal. Then, the same five participants were asked to complete 11 closed-ended questions regarding the age, weight, height, body type, hair colour, length and type, race, facial hair, glasses and eye colour of the criminals.

Based on the written responses from these five participants, the experimenter searched through a database of photographs[1] for individuals that met the descriptions from the five participants. For each ERP-lineup, six photographs of individuals were selected to act as fillers. The six photographs for each lineup were then shown to the five participants that made the written descriptions. For each ERP-lineup, the participants were asked *which of the six photographs looks the most like the criminal?* The photograph that was most agreed upon by the five participants was selected as the *innocent suspect* to replace the criminal in the criminal absent lineups. Therefore, the criminal-present (CP) lineups contained six photographs, five fillers and the criminal and the criminal-absent (CA) lineups contained six photographs, five fillers and the innocent suspect.

Mock witness experiment

Participants, stimuli and procedures

Once the lineup members for each of the four ERP-lineups were selected, an additional 20 participants were provided with one of the five written descriptions of the criminal (counterbalanced across participants). Without ever actually seeing the crime videos, each participant was shown four lineups, two of which where CP and two of which were CA lineups. For each lineup, the six photographs were sequentially shown to the participants. After each photograph, participants were asked to indicate if the photograph was of the criminal or not based on the written descriptions that they were given. They were also asked to rate how certain they were that the photograph was of the criminal on a scale of 0-10 (refer to the certainty ratings

---

[1] Photograph database provided by Steven Smith.

section of Experiment 1 methods) . Participants were also informed that *the criminal might or might not be present in the lineup.*

Results

Eight one-way ANOVAs were conducted with PHOTO (F1, F2, F3, F4, F5 and C/IS) as the independent variable. Separate ANOVAs were conducted for each of the four CP lineups and for each of the four CA lineups, with the certainty ratings for each photograph as the dependant variable. The results indicated that none of the certainty ratings for each photograph was not significantly different from each other for any of the lineups. In other words, no one lineup member (including the criminal and innocent suspect) in any of the lineups was rated with a significantly higher certainty rating than the other lineup members ($p < 0.05$ for all eight ANOVAs).

Discussion

The results from this study demonstrate that the fillers for each of the lineups used in this study (for both CP and CA lineups) share overlapping features with the criminal and were non-biased based on the mock witness procedure.

References

Ashcroft, J., Daniels, D., & Hart, S. (2003). *Eyewitness evidence: a trainer's manual for law enforcement.* US Department of Justice, Office of Justice Programs. National Institute of Justice.

Doob, A.N., & Kirshenbaum, H. (1973). Bias in police lineups - partial remembering. *Journal of Police Science and Administration, 1*, 287-293.

Reno, J., Fisher, R.C., Robinson, L., Brennan, N., & Travis, J. (1999). *Eyewitness Evidence: a guide for law enforcement.* U.S. Department of Justice, Office of Justice Programs. National Institute of Justice.

Wells, G.L., Small, M., Penrod, S. J., Malpass, R. S., Fulero, S. M., & Brimacombe, C. A. E. (1998). Eyewitness identification procedures: Recommendations for lineups and photospreads. *Law and Human Behavior, 22*, 603-647.

Lineup selection and mock witness forms

Written description of the criminal form

In the space provided, please provide a description of the criminal shown in the film:
Video #1 ____________________________

____

____

____

____

____

____

Video #2 ___________________________

____

____

____

____

____

____

Video #3 __________________________

____

____

____

____

____

Video #4 __________________________

____

____

____

Close-ended questionnaire for the description of the criminal

1. What is your best estimate of the perpetrator's AGE? ____________
2. What is your best estimate of the perpetrator's HEIGHT? __________
3. What is your best estimate of the perpetrator's WEIGHT? __________

Circle the appropriate answer:

4. How would you describe the perpetrator's BODY TYPE?
    a. thin
    b. average
    c. heavy
5. How would you describe the perpetrator's HAIR COLOUR?
    a. Blonde
    b. Light brown
    c. Dark brown
    d. Black
    e. Other _____________
6. How would you describe the perpetrator's HAIR TYPE?
    a. Straight
    b. Curly
    c. Wavy
    d. Stringy
    e. Other _________________
7. How would you describe the perpetrator's HAIR LENGTH
    a. Long (past shoulders)
    b. Long (shoulder length)
    c. Medium (part way up neck)
    d. Short (ears exposed)
8. How would you describe the perpetrator's RACE?
    a. White
    b. Oriental
    c. Black
    d. Other ___________
    e. Don't know
9. How would you describe the perpetrator's FACIAL HAIR
    a. None
    b. Moustache
    c. Beard
10. How would you describe the perpetrator's EYE COLOUR
    a. Brown
    b. Blue
    c. Green
    d. Other _____________
11. How would you describe the perpetrator's GLASSES
    a. None
    b. Metal rims
    c. Plastic rims

Appendix D: Forms used in Experiments 1, 2a and 2b

ERP Health Questionnaire

Subject Code: ____________ Sex: M F DOB (d/m/y) ___/___/_____
Occupation:
Highest grade achieved:
Language(s) (1st, fluency of others):
Hand Preference: see Edinburgh Handedness Inventory

Do any of these conditions apply to you:

Stroke

Major surgery (involving generalized anesthetics)

Head injuries

Loss of consciousness

Seizures

Fainting spells, dizziness

Loss of hearing

Perceptual problems (e.g. colour blindness)

Visual disturbances (e.g. cloudy vision, inability to see certain parts of the visual field, lazy eye)

Temporary blindness

Paralysis

Co-ordination problems

Psychological Disorder

Thought disorder

Hallucinations

Learning Disabilities

Drinking/drug dependency

Skin Conditions (e.g psoriasis, eczema)

If you wear glasses did you bring them with you? YES / NO

Experiment 1 rating sheet

Participant Code: ______________ **Date:** __________________

Level of Certainty:

0 —|—|—|—|— **5** —|—|—|—|— **10**

Certain not the criminal	**Unsure if criminal or not**	**Certain is the criminal**

Condition #1
Line-up: ____ Peter _____ Simon

Rate each photo in accordance to how strongly you feel they are the criminal:

#1: __________

#2: __________

#3: __________

#4: __________

#5: __________

#6: __________

Condition #2
Line-up: ____ Peter _____ Simon

#1: __________

#2: __________

#3: __________

#4: __________

#5: __________

#6: __________

Strategy:

Experiments 2a and 2b Rating Sheet

Participant Code: _________ Date: _______________

Level of Certainty:

0 ——+——+——+——+—— 5 ——+——+——+——+—— 10

(certain not criminal) | (unsure if criminal or not) | (certain is the criminal)

Rate each photo in accordance to how strongly you feel they are the criminal:

Line up: _________

1. ______
2. ______
3. ______
4. ______
5. ______
6. ______

Line up: _________

1. ______
2. ______
3. ______
4. ______
5. ______
6. ______

Line up: _________

1. ______
2. ______
3. ______
4. ______
5. ______
6. ______

Line up: _________

1. ______
2. ______
3. ______
4. ______
5. ______
6. ______

Following each line-up rating, Ask:
You selected #___ as the highest value. How confident (0-100%) are you that you selected the correct criminal?

Confidence

1. _______
2. _______
3. _______
4. _______